

Durham Research Online

Deposited in DRO:

24 April 2019

Version of attached file:

Published Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Haehl, Felix M. (2015) 'Comments on universal properties of entanglement entropy and bulk reconstruction.', Journal of high energy physics., 2015 (10). p. 159.

Further information on publisher's website:

[https://doi.org/10.1007/JHEP10\(2015\)159](https://doi.org/10.1007/JHEP10(2015)159)

Publisher's copyright statement:

This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Comments on universal properties of entanglement entropy and bulk reconstruction

Felix M. Haehl

*Centre for Particle Theory & Department of Mathematical Sciences, Science Laboratories,
South Road, Durham DH1 3LE, U.K.*

E-mail: f.m.haehl@gmail.com

ABSTRACT: Entanglement entropy of holographic CFTs is expected to play a crucial role in the reconstruction of semiclassical bulk gravity. We consider the entanglement entropy of spherical regions of vacuum, which is known to contain universal contributions. After perturbing the CFT with a relevant scalar operator, also the first order change of this quantity gives a universal term which only depends on a discrete set of basic CFT parameters. We show that in gravity this statement corresponds to the uniqueness of the ghost-free graviton propagator on an AdS background geometry. While the gravitational dynamics in this context contains little information about the structure of the bulk theory, there is a discrete set of dimensionless parameters of the theory which determines the entanglement entropy. We argue that for every (not necessarily holographic) CFT, any reasonable gravity model can be used to compute this particular entanglement entropy. We elucidate how this statement is consistent with AdS/CFT and also give various generalizations. On the one hand this illustrates the remarkable usefulness of geometric concepts for understanding entanglement in general CFTs. On the other hand, it provides hints as to what entanglement data can be expected to provide enough information to distinguish, e.g., bulk theories with different higher curvature couplings.

KEYWORDS: AdS-CFT Correspondence, Classical Theories of Gravity, Gauge-gravity correspondence

ARXIV EPRINT: [1508.00766](https://arxiv.org/abs/1508.00766)

Contents

1	Introduction	1
2	Universality of entanglement entropy of spherical regions	4
2.1	Setup	4
2.2	CFT explanation for universality	6
2.3	Gravitational explanation for universality	8
2.3.1	Holographic parameter matching	8
2.3.2	Construction of the most general gravitational model	10
2.3.3	Second order Wald entropy is independent of graviton normalization	16
3	Example: Gauss-Bonnet theory	18
3.1	Two gravity models	18
3.2	Parameter matching	19
4	Discussion	21
A	Generalization to bulk theories with higher order equations of motion	25
A.1	Example: general curvature squared theories	25
A.2	General argument	27
B	Perturbations of Wald entropy: background field expansion	27
C	First law of entanglement	30
D	Second order gravitational perturbation theory	32

1 Introduction

It is by now well known that some conformal field theories (CFTs) admit a dual description in terms of semiclassical gravity in one higher dimension [1]. But it is so far not clear what are the precise CFT data that one needs to know in order to reconstruct various specific features of the bulk theory. A full knowledge of the CFT should be exactly equivalent to full knowledge of bulk quantum gravity, the challenge being to find the precise dictionary. We wish to focus on the question which subset of CFT parameters one needs to have access to in order to reconstruct just a classical bulk geometry and its linearized dynamics. Once this question is answered in terms of fundamental CFT parameters, the practical question is which CFT observables are sensitive to these parameters in a useful way.

One hint is the fact that a local bulk metric should be represented in the boundary CFT in a non-local fashion. An important CFT feature which does have non-local properties is the entangled structure of the underlying quantum state. It has therefore been proposed that knowing enough about the entanglement structure of the CFT state is sufficient to determine the gravitational dynamics of the bulk. In particular, entanglement entropy of spatial regions in the CFT appears to be an illuminating quantity to study in this context, cf., [2–6]. The basic rationale for this idea is that entanglement entropy of a given spatial region \mathcal{A} can be computed in semiclassical holography as the area of a certain extremal surface [7–9] which, of course, encodes information about the metric. By varying this entanglement entropy (for instance by changing \mathcal{A} , or by varying the underlying quantum state, or by deforming the CFT spectrum), the corresponding change of the minimal surface encodes information about the dynamics of the bulk metric. Making this idea precise and general would be significant progress towards our understanding of the holographic nature of quantum states and gravity.

We wish to carefully distinguish differently strong versions of the expectation that entanglement encodes bulk geometry. The weakest version of the “*entanglement is geometry*” statement simply says that the most efficient way of calculating entanglement entropy is by using geometric methods. For holographic CFTs the paradigmatic manifestation of this statement is the Ryu-Takayanagi conjecture [7, 8]. But even for non-holographic CFTs there is evidence that geometric concepts sometimes provide the most natural and efficient way of computing entanglement entropy. The first such statement is the realization that vacuum entanglement entropy of spherical regions can be conformally mapped to thermal entropy which sometimes has a natural interpretation in terms of black hole thermodynamics [10]. Refining this idea, Faulkner [11] revisited the problem of computing the response of vacuum entanglement entropy of spheres to a deformation of the CFT by a relevant scalar operator [12–17]. For this computation, nothing is assumed about the existence or the nature of a holographic dual of the CFT. As a CFT calculation this is a formidable problem. Intriguingly, the most convenient way of repackaging this calculation turns out to be in terms of an auxiliary gravitational problem, where the entanglement entropy is represented as a minimal surface area in cosmological Einstein gravity responding to a scalar field perturbation. This is a powerful result, in particular in face of the fact that entanglement entropy is notoriously hard to compute otherwise. One of the goals of this paper is to elucidate from a gravitational point of view how this universality comes about and to reconcile these statements with AdS/CFT duality (where bulk descriptions generically have a structure which is not necessarily given by just Einstein gravity).

A stronger form of the “entanglement is geometry” conjecture is the idea that in holographic CFTs the detailed bulk dynamics can be extracted entirely from knowledge of entanglement entropy for a suitable set of regions \mathcal{A} . Even in the semiclassical regime, this is a strong assertion because the detailed bulk dynamics may involve not only Einstein gravity, but also higher curvature corrections [18–24]. The imprint of higher curvature corrections on equations of motion should then be reconstructible from CFT entanglement entropies. See [4, 6, 25] for recent progress towards deriving linearized Einstein equations from a first law of entanglement [5, 26–28]. It will be a topic of this paper to illuminate the

way in which CFT entanglement does or does not discriminate between different higher curvature interactions in the bulk.

Let us compare these two versions of the “entanglement is geometry” statement. If spherical region entanglement entropy in deformations of any arbitrary CFT can be computed using linearized Einstein gravity [11], it follows that this quantity is certainly not sufficient to distinguish even CFTs with a semiclassical gravity dual from those without; let alone distinguish a bulk theory governed by Einstein gravity from any other higher derivative theory with the same spectrum of low energy excitations. Therefore, for computing this particular type of entanglement entropy, geometry is always an excellent tool; but in order to make sure that this geometry is actually the lowest order manifestation of a semiclassical AdS/CFT duality, more input is required. Of course, the bulk reconstruction from CFT data proceeds gradually: the more CFT parameters the entanglement entropies considered are sensitive to, the more details of the bulk dynamics are a-priori expected to be derivable from it.

In this note we wish to investigate how to find entanglement entropies which one can expect to be sufficient to reconstruct various bulk features such as linearized dynamics, higher curvature couplings etc. In the context of spherical region entanglement entropy and scalar operator deformations, we will give a detailed analysis both from a CFT point of view and from the bulk perspective. As we will see, at the first relevant order in the perturbation, the entanglement entropy is only sensitive to few correlation functions which have a universal form (up to normalization) thanks to conformal invariance. From the bulk perspective the analogous statement is that the backreaction of a maximally symmetric geometry to the presence of a scalar probe is universal for any acceptable theory of gravity.¹ A careful comparison reveals that this entanglement data is not sufficient to conclude anything about the presence of higher curvature terms in the bulk Lagrangian. For the purpose of bulk reconstruction we will outline how to gain sensitivity to such features, but will also demonstrate explicitly that the existence of some higher curvature interactions is much easier to conclude than the precise structure of the latter.

The structure of this note is as follows. In section 2 we first review the context of our discussion and then give CFT and gravity arguments to demonstrate universal properties of entanglement entropy of spherically symmetric regions. In section 3 we illustrate our abstract discussion with the example of curvature squared theories of gravity and Gauss-Bonnet theory in particular. Finally, section 4 contains a discussion and relates to recent developments in this context. We generalize our discussion to a wider class of higher curvature theories in section A, allowing also for fourth order equations of motion. In section B we perform the most general background field expansion for gravity at second order in the metric perturbation, thus uncovering the universality of Gauss-Bonnet theory within this class of theories. Some facts about the first law of entanglement are reviewed in section C. Various useful formulae can be found in section D.

¹By an “acceptable theory of gravity” we mean any semiclassical gravitational theory with the correct spectrum of excitations (in particular without ghosts). See however section A.

2 Universality of entanglement entropy of spherical regions

In this section we analyze universal features both of CFTs and of gravitational theories, which explain universal properties of entanglement entropy of ball shaped spatial regions in CFTs deformed by a scalar operator. We will then compare these using holography and draw conclusions for bulk reconstruction from entanglement.

2.1 Setup

Consider a holographic CFT_d in the vacuum state and on Minkowski spacetime. Assume further that the theory has a classical gravity dual described by Einstein theory and consider a ball shaped spatial region \mathcal{A} of radius R . The reduced density matrix associated with \mathcal{A} is obtained by tracing out the degrees of freedom outside of \mathcal{A} :

$$\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{A}^c} |0\rangle\langle 0|. \quad (2.1)$$

Without loss of generality, we will assume $\rho_{\mathcal{A}}$ is normalized, i.e., $\text{Tr} \rho_{\mathcal{A}} = 1$. The entanglement entropy of \mathcal{A} is defined as the von Neumann entropy of $\rho_{\mathcal{A}}$ and can be computed in the bulk as the area of the minimal surface ending on the boundary of the ball [7]:

$$S_{EE}^{(0)}(\mathcal{A}) = -\text{Tr}(\rho_{\mathcal{A}} \log \rho_{\mathcal{A}}) = \frac{\text{Vol}(\mathcal{E}_{\min})}{4G_N}, \quad (2.2)$$

where \mathcal{E}_{\min} is the corresponding bulk minimal surface and G_N denotes Newton's constant in $d+1$ dimensions. Since the region \mathcal{A} is spherical, its causal development can be mapped by a conformal transformation to a direct product of time with a maximally symmetric hyperbolic space [10]. This transformation further maps the CFT vacuum state to a thermal state. The extremal surface is no longer anchored at the boundary but instead wraps the horizon of a hyperbolic black hole. The computation of entanglement entropy of \mathcal{A} thus reduces to a calculation of the horizon entropy of a hyperbolic black hole at temperature $T = 1/(2\pi R)$:

$$S_{EE}^{(0)}(\mathcal{A}) = S_{\text{thermal}}(\mathbb{H}_{d-1}) = \frac{\text{Vol}(\mathbb{H}_{d-1})}{4G_N}, \quad (2.3)$$

where \mathbb{H}_{d-1} is the hyperbolic horizon slice. For notational simplicity, we will from now on work with Planck units, i.e., $\ell_{\text{P}}^{d-1} = 8\pi G_N$.

Now let us ask what changes if the bulk gravitational theory is not just Einstein gravity, but also contains some higher curvature corrections. For concreteness, consider a $d+1$ -dimensional bulk theory of the form

$$I = \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \mathcal{L} \equiv \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \left(R + \frac{d(d-1)}{\ell^2} + \mathcal{L}_{h.d.} \right), \quad (2.4)$$

where $\mathcal{L}_{h.d.} \equiv \mathcal{L}_{h.d.}(g_{ab}, R_{abcd}, \nabla_e R_{abcd}, \dots)$ encodes higher curvature corrections. The cosmological length scale ℓ is related to the AdS scale ℓ_{AdS} ; e.g., in Einstein gravity, $\ell = \ell_{\text{AdS}}$.

The thermal entropy (and thus the entanglement entropy) is now given by the Wald entropy [29] of the hyperbolic horizon [19, 30, 31]:

$$\begin{aligned} S_{EE}^{(0)}(\mathcal{A}) &= S_{\text{Wald}}(\mathbb{H}_{d-1}) = \frac{-2\pi}{2\ell_{\text{P}}^{d-1}} \int_{\mathbb{H}_{d-1}} \sqrt{\gamma} \left(-2 + \frac{\delta \mathcal{L}_{h.d.}}{\delta R_{abcd}} n_{ab} n_{cd} \right) \\ &= \frac{2\pi}{\pi^{d/2}} \Gamma(d/2) a_d^* \frac{\text{Vol}(\mathbb{H}_{d-1})}{\ell_{\text{AdS}}^{d-1}}, \end{aligned} \quad (2.5)$$

where n_{ab} denotes the binormal to the hyperbolic horizon \mathbb{H}_{d-1} normalized as $n^{ab}n_{ab} = -2$ and γ is the induced metric. The constant a_d^* depends on the details of the gravitational theory. For example, in Einstein gravity $\mathcal{L}_{h.d.} = 0$ and we find by comparison with (2.3)

$$(a_d^*)_{\text{Einstein}} = \frac{\pi^{d/2}}{\Gamma(d/2)} \frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}}. \quad (2.6)$$

In fact, since the Wald entropy density is constant on maximally symmetric AdS_{d+1} , we can immediately read off the general expression for the constant a_d^* :

$$a_d^* = \frac{\pi^{d/2}}{\Gamma(d/2)} \frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \left(1 - \frac{1}{2} \frac{\delta \mathcal{L}_{h.d.}}{\delta R_{abcd}} n_{ab} n_{cd} \Big|_{\text{AdS}} \right) = -\frac{\pi^{d/2} \ell_{\text{AdS}}^{d+1}}{d \Gamma(d/2)} \frac{1}{2\ell_{\text{P}}^{d-1}} \mathcal{L}|_{\text{AdS}}. \quad (2.7)$$

Here, the second step is a simple consequence of the existence of an AdS_{d+1} background solution and the maximal symmetry of the geometry (see, e.g., [30] for a detailed derivation).² The important feature of (2.5) is that the entanglement entropy is always proportional to the horizon area $\text{Vol}(\mathbb{H}_{d-1})$ (or equivalently to $\text{Vol}(\mathcal{E}_{\text{min}})$), irrespective of the details of the gravitational theory. In particular, just knowing $S_{EE}^{(0)}(\mathcal{A})$ does not give any insight as to what higher derivative corrections the bulk dual of the given CFT may entail: any higher derivative corrections to Einstein gravity will only change the overall normalization but not the volume scaling in (2.5). But any such change of a_d^* due to higher curvature corrections can be equivalently interpreted as simply a renormalization of the ratio $\ell_{\text{AdS}}/\ell_{\text{P}}$ within Einstein gravity. The precise renormalization factor is given by the bracket in (2.7). The effect of higher derivative terms is hence not visible for its only imprint is to renormalize a dimensionless parameter of the theory.

Note that the above expressions are divergent and the matchings performed should hence be done more carefully. For example, the volume factor in (2.3) should be regulated such that its leading divergent piece scales with the area of the sphere \mathcal{A} . This divergence depends on the cut-off, but in a consistent way such that the CFT short distance cut-off (regulating the UV divergence of entanglement entropy near the boundary of \mathcal{A}) matches the bulk cut-off (regulating the divergent area of \mathbb{H}_{d-1}) [8]. More importantly, the expression (2.3) also contains a universal piece which is constant in even dimensions and logarithmically divergent in odd dimensions. For this reason, the above matching conditions strictly speaking refer to the universal piece, which takes the form [8, 31]

$$S_{EE,\text{univ}}^{(0)}(\mathcal{A}) = \begin{cases} -(-1)^{d/2} 4 a_d^* \log(2R/\delta) & (d \text{ even}) \\ (-1)^{(d-1)/2} 2\pi a_d^* & (d \text{ odd}) \end{cases} \quad (2.8)$$

²The first term in the bracket of (2.7) comes from the Einstein contribution in the action (2.4). In a theory $\mathcal{L} = d(d-1)/\ell^2 + \mathcal{L}_{h.d.}$ without pure Ricci scalar, this term would disappear.

where δ is the UV cut-off scale in the CFT. It is only for spherical regions \mathcal{A} that our arguments above give the correct result without explicitly dealing with these additional subtleties, see e.g. [19].

From a CFT point of view, a_d^* is a fundamental parameter. As we have just seen, it characterizes the normalization of the universal part of vacuum entanglement entropy. It is a remarkable fact about entanglement entropy that in even dimensions a_d^* coincides with the coefficient a of the Euler density in the conformal trace anomaly [30–32]. In odd dimensions, a_d^* is proportional to $\log Z_{S^d}$, the sphere partition function of the CFT [10].

For a detailed reconstruction of bulk dynamics from entanglement, one would like to probe the particular structure of the gravitational action. In this note we wish to study deformations of the above setup and ask what CFT quantities one can use in order to gain sensitivity to features which are peculiar to higher curvature terms in the gravity dual. Clearly the vacuum entanglement entropy of balls \mathcal{A} will not be enough by the above argument. Consider therefore a deformation of the CFT by some relevant scalar operator \mathcal{O} , uniformly coupled to the CFT with a coupling λ . Given the arguments above, it is clear what bulk computation we should do in order to learn about the change in $S_{EE}(\mathcal{A})$: add to the gravitational problem a scalar field dual to \mathcal{O} with suitable boundary conditions, solve the scalar equation of motion in the AdS_{d+1} background and compute the backreaction on the minimal surface area perturbatively in λ . As we will see, the linearized answer is still universal in a way very similar to how (2.5) was universal for the unperturbed problem. Namely, its functional form is completely fixed such that it only depends on a few parameters which encode the influence of higher curvature couplings. These parameters can again be absorbed in a renormalization of dimensionless quantities such that the answer is indistinguishable from what one would have obtained in Einstein gravity minimally coupled to a scalar field. This is the main universality statement studied in this section: *For holographically computing entanglement entropy of spherical regions in CFTs deformed by a relevant scalar operator, any physically acceptable theory of a spin-2 and a spin-0 field with appropriately tuned values of the couplings gives the same result as cosmological Einstein gravity minimally coupled to a scalar. In particular, this quantity does not distinguish between different higher curvature theories of gravity.* In the rest of this section, we will explain these statements from the CFT and from the gravity side.

2.2 CFT explanation for universality

We start by reviewing from a CFT point of view the ingredients involved in computing entanglement entropy of spherical regions in perturbed states. Our discussion closely follows [11, 16, 17].

Consider a CFT on d -dimensional Euclidean flat space and perform the following deformation of the CFT action $I_{CFT}^{(0)}$ by a relevant scalar operator \mathcal{O} of dimension $\Delta < d$:

$$I_{CFT} = I_{CFT}^{(0)} + \lambda \int d^d x \mathcal{O}(x). \quad (2.9)$$

We can compute the perturbative changes of the entanglement entropy $S_{EE}(\mathcal{A})$ by writing the reduced density matrix as a path integral. This path integral lives on a manifold which

is the Euclidean spacetime with a cut along $\partial\mathcal{A}$. The matrix element corresponding to some field configurations ϕ_+ and ϕ_- on the two sides of the cut is

$$\begin{aligned}\langle\phi_-|\rho_{\mathcal{A}}|\phi_+\rangle &= \int_{\substack{\phi(\partial\mathcal{A}_+)=\phi_+ \\ \phi(\partial\mathcal{A}_-)=\phi_-}} [\mathcal{D}\phi] e^{-I_{CFT}} \\ &= \int_{\substack{\phi(\partial\mathcal{A}_+)=\phi_+ \\ \phi(\partial\mathcal{A}_-)=\phi_-}} [\mathcal{D}\phi] e^{-I_{CFT}^{(0)}} \left(1 - \lambda \int d^d x \mathcal{O}(x) + \frac{\lambda^2}{2} \iint d^d x d^d y \mathcal{O}(x) \mathcal{O}(y) + \dots \right).\end{aligned}\quad (2.10)$$

Now define a modular Hamiltonian $H_{\mathcal{A}}$ via $\rho_{\mathcal{A}} = e^{-H_{\mathcal{A}}}$. We can then use the identity $S_{EE}(\mathcal{A}) = \text{Tr}(\rho_{\mathcal{A}} H_{\mathcal{A}})$ to compute the change in the entanglement to any desired order in λ :³

$$\begin{aligned}S_{EE}(\mathcal{A}) &= S_{EE}^{(0)}(\mathcal{A}) + \lambda \left[\frac{\partial S_{EE}(\mathcal{A})}{\partial \lambda} \right]_{\lambda=0} + \frac{\lambda^2}{2} \left[\frac{\partial^2 S_{EE}(\mathcal{A})}{\partial \lambda^2} \right]_{\lambda=0} + \dots \\ &= S_{EE}^{(0)}(\mathcal{A}) + \lambda \left[-\langle \mathcal{O} H_{\mathcal{A}} \rangle + \left\langle \frac{\partial H_{\mathcal{A}}}{\partial \lambda} \right\rangle \right]_{\lambda=0} + \frac{\lambda^2}{2} \left[\langle \mathcal{O} \mathcal{O} H_{\mathcal{A}} \rangle - \langle \mathcal{O} \frac{\partial H_{\mathcal{A}}}{\partial \lambda} \rangle + \left\langle \frac{\partial^2 H_{\mathcal{A}}}{\partial^2 \lambda} \right\rangle \right]_{\lambda=0} + \dots \\ &= S_{EE}^{(0)}(\mathcal{A}) - \lambda \langle \mathcal{O} H_{\mathcal{A}}^{(0)} \rangle + \frac{\lambda^2}{2} \left(\langle \mathcal{O} \mathcal{O} H_{\mathcal{A}}^{(0)} \rangle - \langle \mathcal{O} \frac{\partial H_{\mathcal{A}}}{\partial \lambda} \rangle \right) + \dots\end{aligned}\quad (2.11)$$

where $S_{EE}^{(0)}(\mathcal{A})$ is universal up to an overall normalization a_d^* as explained in section 2.1, and similarly $H_{\mathcal{A}}^{(0)}$ is the unperturbed modular Hamiltonian. In the last step we used the fact that $\rho_{\mathcal{A}}$ is normalized which implies $\langle \frac{\partial H_{\mathcal{A}}}{\partial \lambda} \rangle = 0$. We can now use that \mathcal{A} is a spherical region, which brings about another important simplification: the modular Hamiltonian can be traded for an integrated stress tensor, since for spherical \mathcal{A} with radius R centered at the origin, one finds

$$H_{\mathcal{A}} = 2\pi \int_{\mathcal{A}} d^{d-1}x \frac{R^2 - \vec{x}^2}{2R} T_{00}(x), \quad (2.12)$$

for \mathcal{A} lying on the $x^0 = 0$ slice. In our problem, correlation functions involving the modular Hamiltonian can therefore be regarded as correlation functions involving the conformal stress tensor. But for stress tensor correlation functions, we can use well-known features of CFTs: on general grounds, two-point correlations between a stress tensor and a primary operator are constrained by conformal invariance and conformal Ward identities [33]. In fact, independent of the details of our problem, conformal invariance of a theory in a flat spacetime enforces $\langle \mathcal{O} T_{\mu\nu} \rangle = 0$. Therefore the $\mathcal{O}(\lambda)$ perturbation in (2.11) vanishes.⁴

Furthermore, the perturbation of $T_{\mu\nu}$ is uniform in \mathcal{O} and we have $\frac{\partial T_{\mu\nu}}{\partial \lambda} = -\delta_{\mu\nu} \mathcal{O}$. To see this, recall that for a Euclidean flat boundary, we have

$$T_{\mu\nu}(x) = \frac{2}{\sqrt{g}} \frac{\delta I_{CFT}}{\delta g^{\mu\nu}(x)} = T_{\mu\nu}^{(0)}(x) - \delta_{\mu\nu} \lambda \mathcal{O}(x). \quad (2.13)$$

The entanglement entropy in the deformed theory, (2.11), hence takes a generic form [17]:

$$S_{EE}(\mathcal{A}) = S_{EE}^{(0)}(\mathcal{A}) + \frac{\lambda^2}{2} (\langle \mathcal{O} \mathcal{O} H_{\mathcal{A}}^{(0)} \rangle - \langle \mathcal{O} \mathcal{O} \rangle) + \dots \quad (2.14)$$

³Operators without argument are understood to be integrated, e.g., $\mathcal{O} \equiv \int d^d x \mathcal{O}(x)$.

⁴We will use indices a, b, \dots for bulk coordinates, μ, ν, \dots for boundary directions, and α, β, \dots for intrinsic coordinates on the minimal surface \mathcal{E}_{min} .

As we can see, this is only sensitive to the correlators $\langle \mathcal{O} \mathcal{O} T_{\mu\nu} \rangle$ and $\langle \mathcal{O} \mathcal{O} \rangle$. These, however, are known to be universal functions in any conformal field theory which only depend on one overall normalization constant:

$$\langle \mathcal{O}(x_1) \mathcal{O}(x_2) \rangle = \frac{C_{\mathcal{O}}}{(x_1 - x_2)^{2\Delta}} \quad (2.15)$$

and a similar universal function for $\langle \mathcal{O}(x_1) \mathcal{O}(x_2) T_{\mu\nu}(x_3) \rangle$ whose normalization is also determined solely by $C_{\mathcal{O}}$ [33, 34]. The entanglement entropy of \mathcal{A} is hence completely universal up to $\mathcal{O}(\lambda^2)$, only depending on the fundamental constants $\{a_d^*, C_{\mathcal{O}}\}$ and the operator dimension Δ .

The divergent structure of the expression (2.14) has been computed in [17] just using standard Euclidean CFT techniques. However, we must now face an additional subtlety: the analysis of [11], which treats the problem at hand using the replica trick and carefully analyzes the analytic structure of the resulting partition functions, leads to a different result which involves a non-trivial finite piece. Nevertheless the basic lesson of this more involved analysis using the replica trick turns out to be still the same; namely the $\mathcal{O}(\lambda^2)$ perturbation of entanglement entropy is determined solely by correlators $\langle \mathcal{O} \mathcal{O} \rangle$ and $\langle \mathcal{O} \mathcal{O} T_{\mu\nu} \rangle$. To be precise, the finite contribution to the correction at $\mathcal{O}(\lambda^2)$ reads [11, 14, 15]:⁵

$$S_{EE}(\mathcal{A}) = S_{EE}^{(0)}(\mathcal{A}) - \lambda^2 R^{2(d-\Delta)} \frac{\pi^{\frac{d+1}{2}} (d-\Delta) \Gamma(1 + \frac{d}{2} - \Delta)}{2 \Gamma(\frac{3}{2} + d - \Delta)} + \dots, \quad (2.16)$$

where R is the radius of the ball \mathcal{A} . In this expression the normalization $C_{\mathcal{O}}$ has been fixed in a way that will prove convenient when comparing with holographic calculations. However, we will not need the explicit form of $S_{EE}(\mathcal{A})$ for the following analysis.

2.3 Gravitational explanation for universality

Having seen that the second order computation of perturbed entanglement entropy of spherical regions is only sensitive to $\langle \mathcal{O} \mathcal{O} \rangle$ and $\langle \mathcal{O} \mathcal{O} T_{\mu\nu} \rangle$ correlators (which are universal up to one normalization constant), we now want to explain what this means in gravity.

2.3.1 Holographic parameter matching

A major hint for understanding the entanglement universality from a holographic point of view is the result of [11], where it is shown that for the CFT deformation characterized by a coupling λ and an operator dimension Δ as described in section 2.2, one can calculate the entanglement entropy at $\mathcal{O}(\lambda^2)$ by solving an auxiliary gravitational problem described by the action

$$I_{\text{aux}} = \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \left[R + \frac{d(d-1)}{\ell^2} - \frac{\lambda_{\phi}}{2} ((\nabla\phi)^2 + m^2 \phi^2) \right]. \quad (2.17)$$

The AdS_{d+1} background geometry \bar{g}_{ab} has a scale which coincides with the cosmological scale: $\ell_{\text{AdS}}^2 = \ell^2$. Further, λ_{ϕ} parametrizes the normalization of the scalar action. The

⁵This is assuming that \mathcal{O} is a relevant operator with $\Delta \neq d/2$. See [14, 15] for a discussion of the case $\Delta = d/2$ in holographic theories.

prescription is to solve first the scalar wave equation in the background \bar{g}_{ab} and then compute its backreaction on the metric at next order in perturbation theory to get the change of the area of the Ryu-Takayanagi minimal surface. For this task we are therefore not interested in the full action (2.17), but only in the linearized equations of motion due to a scalar probe in AdS_{d+1} background. In the following all objects with a bar (e.g., $\bar{\nabla}_a$) refer to the background metric \bar{g}_{ab} .

We will now comment on the number of free parameters in this problem and how to determine them, given that they should induce the CFT parameters at the boundary in a way consistent with AdS/CFT. A priori there are three dimensionless quantities $\{(\ell_{\text{AdS}}/\ell_{\text{P}}), \ell_{\text{AdS}}^2 m^2, \lambda_\phi\}$ in the action (2.17). The scalar wave equation in AdS, $(\bar{\square} - m^2)\phi = 0$, has two asymptotic solutions with respective falloff behavior z^{Δ_+} and z^{Δ_-} , where the boundary is at $z = 0$ (in Poincaré coordinates such as (C.4)) and

$$\Delta_{\pm} \equiv \frac{d}{2} \pm \sqrt{\frac{d^2}{4} + \ell_{\text{AdS}}^2 m^2}. \quad (2.18)$$

The boundary condition for the scalar is required to be such that it sources an operator with given dimension Δ and coupling λ . By standard AdS/CFT techniques [1], we therefore assume⁶ $\Delta \geq d/2$, tune the scalar mass such that $\Delta_+ = \Delta$ and fix the dominant asymptotics as $\phi(z \rightarrow 0) \sim \lambda z^{\Delta_-}$. The mass of the scalar is also immediately fixed by demanding that it sources an operator with dimension Δ ; the usual prescription for this is to take $-d^2/4 < \ell_{\text{AdS}}^2 m^2 < 0$ such that

$$\ell_{\text{AdS}}^2 m^2 = \Delta(\Delta - d). \quad (2.19)$$

The remaining two parameters of the theory also need to be fixed by matching with the CFT. On the one hand, $\ell_{\text{AdS}}/\ell_{\text{P}}$ makes an appearance in the normalization of the universal scaling of entanglement in the CFT in terms of the generalization of a -central charge, cf., (2.5) and can hence be fixed by demanding (2.6). On the other hand, λ_ϕ parametrizes rescalings of the bulk field ϕ in the same sense in which the dual boundary two-point function $\langle \mathcal{O}\mathcal{O} \rangle$ depends on an overall normalization $C_{\mathcal{O}}$ as in (2.15). Since ϕ has already been equipped with boundary conditions which lead to the correct coupling λ for \mathcal{O} , we can follow standard techniques to work out the normalization λ_ϕ in terms of $C_{\mathcal{O}}$. This involves computing the bulk to boundary propagator for ϕ and matching its boundary two-point function with the CFT correlator $\langle \mathcal{O}\mathcal{O} \rangle$. The result is [1, 13]

$$C_{\mathcal{O}} = \frac{\Delta - d/2}{\pi^{d/2}} \frac{\Gamma(\Delta)}{\Gamma(\Delta - d/2)} \frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \lambda_\phi. \quad (2.20)$$

We stress that, as shown in [11], the auxiliary gravitational problem just described can be used for computing perturbed entanglement entropy of ball shaped regions in *any* CFT. Given the universality of the correlators $\langle \mathcal{O}\mathcal{O} \rangle$ and $\langle \mathcal{O}\mathcal{O}T_{\mu\nu} \rangle$ in all CFTs (including

⁶The assumption $\Delta \geq d/2$ corresponds to performing so-called standard quantization. Note that $(d - 2)/2 \leq \Delta \leq d/2$ would still be compatible with field theory unitarity. The latter regime can be treated by choosing $\Delta_- = \Delta$. See [13] for a discussion of this case.

those which are dual to cosmological Einstein gravity), this result is not very surprising. It is, however, quite remarkable that the most natural way of rewriting the CFT problem is in terms of the auxiliary gravitational system (2.17). As we shall see, there are many other (higher curvature) models of gravity which could equally well serve as the auxiliary system for this problem. Turning this logic around, all these different gravitational theories are indistinguishable if only perturbed entanglement data of spheres is measured. We now turn to illuminating this point from the gravity side.

2.3.2 Construction of the most general gravitational model

We now want to make more precise the statement that all physically acceptable gravity models are equivalent for the computation at hand. Our approach is to construct a large class of auxiliary gravitational systems which make the same predictions as (2.17) in this respect. Let us take the CFT data as given and try to construct a holographic description in terms of semiclassical gravity ‘bottom up’. The first step of such a program could be the question: what are the minimal bulk ingredients if they are supposed to compute entanglement entropy of spheres in vacuum and after turning on the perturbation? We will now answer this question by building general linearized bulk actions which satisfy the following conditions:

- The field content is a symmetric spin-2 tensor (the metric) and a scalar field, which source the boundary operators $T_{\mu\nu}$ and \mathcal{O} , respectively. The model gives dynamics to these fields and defines a coupling between them which allows to compute the linearized backreaction h_{ab} on an AdS_{d+1} background \bar{g}_{ab} after perturbing the latter with the scalar probe ϕ .
- There are no ghost degrees of freedom; in particular the linearized equations of motion on AdS_{d+1} are second order in derivatives. (See below for a justification of this assumption and section A for generalizations in the case where the higher curvature terms are perturbatively small and this assumption can be dropped.)
- The action is diffeomorphism invariant (cf., [35, 36]).

We will find that these conditions specify the linearized action uniquely up to normalization constants; see (2.21) for the final result.

Let us briefly explain why the equations of motion should be second order. It is well known that equations of motion which are higher than second order lead to pathological behaviour which we want to exclude from our discussion of theories which are dual to unitary CFTs. In particular, higher order equations of motion generically give rise to further propagating degrees of freedom with negative energy [37–41].⁷ The presence of ghosts in

⁷There are special circumstances where the additional degrees of freedom can be dealt with by using appropriate field redefinitions. For instance, $f(R)$ theories of gravity naively have higher order equations of motion, but are known to be equivalent under field redefinitions to Einstein gravity coupled to a (positive energy) scalar field such that the boundary theory is unitary [42]. The graviton in this case couples not only to the boundary stress tensor, but in addition to some scalar operator. The CFT therefore has some more specific features than we assumed in the beginning. For validating our conclusions in the most general

such theories thus leads to classical instabilities [43] (which would lead to negative norm states and therefore non-unitary quantum theories upon quantization). These features would also be visible in the dual CFT as operators with complex conformal dimension or negative norm. Indeed, via holography the problem of perturbed entanglement entropy is described in the CFT by correlation functions of the stress tensor $T_{\mu\nu}$ and a relevant scalar operator \mathcal{O} . This setup should be described in gravity by nothing more than a spin-2 graviton which sources $T_{\mu\nu}$ and a scalar field to source \mathcal{O} .⁸ If the graviton equations of motion were higher than second order in derivatives, then the additional unwanted ghost modes would couple to other (non-unitary) operators. Therefore, whatever the full gravitational theory is, after linearization of the action around AdS_{d+1} the excitation h_{ab} is bound to appear in such a way as to yield second order equations of motion.

Despite these arguments in favour of two-derivative equations of motion, higher derivative models with higher order equations of motion are regularly studied as toy models for semiclassical corrections to Einstein gravity induced by some consistent truncation of full unitary quantum gravity (which is assumed to be UV complete). Such models can be justified as sensible toy models in a perturbative framework as long as the ghost modes are very heavy and do not go on-shell (which turns out to be the case whenever the higher derivative couplings are small). We explore this setup in section A, where we show that in this context our arguments concerning entanglement entropy still hold. The reason for the latter is that at low energies the ghosts can be ignored and the sector of physical modes is exactly the same as for ghost-free models which we are now going to analyze.

Let us now construct the minimal ingredients for a linearized action for h_{ab} and ϕ , which leads to second order equations of motion. We will argue that the most general action of this type is the same as the linearization of Einstein-scalar theory (2.17) up to various normalizations. For illustration, we can hence start by expanding the metric around an AdS_{d+1} background solution \bar{g}_{ab} as $g_{ab} = \bar{g}_{ab} + h_{ab}$ and linearizing the Einstein action (2.17). Including some convenient normalization factors, this procedure leads to⁹

$$I = \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-\bar{g}} \left\{ -\frac{2d}{\ell_{\text{AdS}}^2} \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{\ell_{\text{P}}^{d-1}}{\ell_{\text{AdS}}^{d-1}} a_d^* + \frac{\lambda_h}{2} h_{ab} \mathcal{G}^{abcd} h_{cd} + \lambda_\phi \left[-\frac{1}{2} ((\bar{\nabla}\phi)^2 + m^2\phi^2) \right] \right. \\ \left. + \frac{\lambda_\phi}{2} h^{ab} \left(-\frac{1}{2} [(\bar{\nabla}\phi)^2 + m^2\phi^2] \bar{g}_{ab} + \bar{\nabla}_a\phi \bar{\nabla}_b\phi \right) + \mathcal{O}(h^3, h^2\phi^2) \right\}, \quad (2.21)$$

circumstances, it would be interesting to explore the most general requirements for a bulk theory to admit a unitary dual CFT. A first step in this direction might be a gauge-fixed analysis where the scalar part of the graviton is set to zero, cf., the transverse traceless gauge employed in section A. This would at least take care of $f(R)$ theories.

⁸Of course, it would be possible for \mathcal{O} to be some other scalar operator, e.g., one corresponding to the square of a scalar field. Since such scenarios would not change our analysis qualitatively, we will assume that \mathcal{O} is simply sourced by ϕ with appropriate boundary condition.

⁹To verify this by hand, the identities in section D are helpful.

where \mathcal{G}_{abcd} is the graviton kinetic operator

$$\begin{aligned} \mathcal{G}_{abcd} = & \frac{1}{2} (\bar{g}_{a(c} \bar{g}_{d)b} - \bar{g}_{ab} \bar{g}_{cd}) \bar{\square} - \bar{\nabla}_{(c} \bar{g}_{d)(a} \bar{\nabla}_{b)} + \frac{1}{2} (\bar{g}_{ab} \bar{\nabla}_c \bar{\nabla}_d + \bar{g}_{cd} \bar{\nabla}_a \bar{\nabla}_b) \\ & + \frac{d}{\ell_{\text{AdS}}^2} \left(\frac{1}{2} \bar{g}_{ab} \bar{g}_{cd} - \bar{g}_{a(c} \bar{g}_{d)b} \right). \end{aligned} \quad (2.22)$$

A few explanations are in order concerning the way we wrote (2.21):

- i) In writing (2.21) we introduced a normalization λ_h in the graviton kinetic term for later convenience. By expanding the Einstein-scalar action (2.17), one finds simply $(\lambda_h)_{\text{Einstein}} = 1$, but we will shortly encounter other (higher derivative) theories with other values of λ_h .
- ii) Similarly, we wrote a constant term proportional to a_d^* . This is also using the benefit of hindsight (in particular (2.7)) to already implement the correct generalization for higher derivative theories. For the present case of Einstein gravity, $(a_d^*)_{\text{Einstein}}$ is just given by (2.6). For more general theories, notice that the constant term in (2.21) is determined by the requirement that switching off h_{ab} and ϕ in the action is equivalent to evaluating the full non-linear Lagrangian on the AdS background. By means of (2.7) we know that the constant thus obtained needs to be a_d^* (with appropriate pre-factors).
- iii) Note also that we did not write the $\mathcal{O}(h)$ term of the action. This term would just define the equations of motion for the AdS_{d+1} background geometry \bar{g}_{ab} . By assuming an expansion of the form $g_{ab} = \bar{g}_{ab} + h_{ab}$, such terms vanish automatically.

The graviton-scalar action (2.21) was obtained by linearizing the Einstein-scalar system (2.17) around a given AdS background with scale ℓ_{AdS} . However, we will now argue that *any* theory which satisfies the basic conditions listed at the beginning of this subsection, takes this form. As we will see, the only freedom is parameterized by the normalizations $\{a_d^*, \lambda_h, \lambda_\phi, m^2\}$. Of these, a_d^* has already been fixed manifestly by using (2.7). Also λ_ϕ and $\ell_{\text{AdS}}^2 m^2$ could be fixed in terms of field theory quantities via (2.19) and (2.20). The graviton normalization, λ_h , will be discussed shortly. This will then demonstrate the universality of the dynamics of perturbing an AdS geometry with a scalar field up to normalizations.

Graviton propagator on AdS_{d+1} . Clearly, the field content in the bulk should involve a spin-0 field ϕ to probe the geometry, whose boundary condition $\phi(z \rightarrow 0) \sim \lambda z^{d-\Delta}$ sources the relevant deformation \mathcal{O} . But before specifying the matter probe in this much detail, let us discuss the general gravitational response to any small perturbation. This is described by a spin-2 field propagating on the background \bar{g}_{ab} . We will write this as $g_{ab} = \bar{g}_{ab} + h_{ab}$, where h_{ab} is the linearized perturbation.

We now want to establish the most general physically acceptable equations of motion for h_{ab} , as described above. It is straightforward to write the most general (up to total derivatives) action at $\mathcal{O}(h^2)$ of the symmetric spin-2 field which involves at most two

derivatives. We can parametrize it as (see, e.g., [44])

$$I_{(h^2)} = \frac{1}{2\ell_P^{d-1}} \int \sqrt{-\bar{g}} \left\{ -\frac{\lambda_h}{4} \bar{\nabla}_a h_{bc} \bar{\nabla}^a h^{bc} + \frac{b_1}{2} \bar{\nabla}_b h^{ac} \bar{\nabla}_a h_c^b - \frac{b_2}{2} \bar{\nabla}^a h \bar{\nabla}^b h_{ab} \right. \\ \left. + \frac{b_3}{4} \bar{\nabla}_a h \bar{\nabla}^a h + \frac{M^2}{2} \left(h^2 - b_4 h_{ab} h^{ab} \right) \right\}, \quad (2.23)$$

where $h \equiv h^{ab} \bar{g}_{ab}$. All raising and lowering of indices as well as covariant derivatives with a bar refer to \bar{g}_{ab} .¹⁰ We can already see that compared to the infinite number of possible higher derivative theories of gravity, the above linearized action is quite restrictive, having only six free parameters $\{\lambda_h, M^2, b_1, b_2, b_3, b_4\}$. In fact the action will be even more constrained once we impose that the full theory should exhibit diffeomorphism invariance. It is a classic result that demanding a diffeomorphism invariant theory beyond the linearized level singles out a unique graviton propagator [45]. In our explicit notation, we can give a simple justification of this statement by considering a spin-2 gauge transformation (aka linearized diffeomorphism) $h_{ab} \rightarrow h_{ab} + 2\partial_{(a} \xi_{b)}$. The $\mathcal{O}(h^2)$ action transforms nontrivially:

$$\delta I_{(h^2)} = \frac{1}{2\ell_P^{d-1}} \int \sqrt{-\bar{g}} \xi^a \left\{ (b_1 - \lambda_h) \bar{\nabla}^b \bar{\square} h_{ab} + (b_1 - b_2) \bar{\nabla}_a \bar{\nabla}_c \bar{\nabla}^a h^{cd} + (b_3 - b_2) \bar{\nabla}_a \bar{\square} h \right. \\ \left. - \left(2M^2 - \frac{d b_2}{\ell_{\text{AdS}}^2} \right) \bar{\nabla}_a h + \left(2b_4 M^2 - \frac{2 d b_1}{\ell_{\text{AdS}}^2} \right) \bar{\nabla}_b h_a^b \right\}, \quad (2.24)$$

where we used partial integration and (D.2). In order for the action to be diffeomorphism invariant, the above expression has to vanish for arbitrary ξ^a . Since all terms are independent, this condition forces upon us five relations among the six parameters:

$$b_1 = b_2 = b_3 = \lambda_h, \quad M^2 = \frac{d}{2\ell_{\text{AdS}}^2} \lambda_h, \quad b_4 = 2. \quad (2.25)$$

Any other choice of parameters would not correspond to the linearization of a diffeomorphism invariant theory. The only free parameter which is left undetermined is the overall scale λ_h of the graviton action. In a full non-linear theory of gravity this is, of course, fixed by the normalization of the action and we will compute some examples in section 3 to illustrate this. For instance, we have already seen that the graviton action (2.23) is exactly the same as the second order term in a perturbative expansion of the Einstein-Hilbert action with cosmological constant scale $\ell = \ell_{\text{AdS}}$ and overall normalization $\lambda_h = 1$.

This construction shows that by just looking at the linearized graviton propagator on AdS of any reasonable diffeomorphism invariant theory of gravity, we will never be able to distinguish the theory from cosmological Einstein gravity. The underlying reason for this is the universality of spin-2 propagation on maximally symmetric backgrounds.

¹⁰Note that the coefficient λ_h of the first term in (2.23) cannot be vanishing as this would mean that no spin-2 degrees of freedom propagate at all. By this we ensure that no gauge choice can eliminate the spin-2 mode. For instance, we could choose to work in a transverse traceless gauge ($\bar{\nabla}_a h^{ab} = 0 = h$), but the first term would still survive as it is the source of the spin-2 component. On a similar note, we need $\lambda_h > 0$ in order for the spin-2 graviton not to be ghost.

Coupling to scalar matter. Having constructed the most general quadratic action for the metric perturbation, we should now in a similar fashion argue that (2.21) couples h_{ab} to the scalar field in the most general way at this order. There are several possible Yukawa couplings and self-interaction terms. However, for consistency of our perturbative scheme, the only terms we are interested in at this order are $\phi\phi$ self-interactions and $h\phi\phi$ couplings. We will now describe the most general actions of these kinds:

- **Scalar field kinetic and mass term:** at lowest order in perturbation theory (the probe limit), we are only concerned with couplings of $\phi\phi$ to the background geometry. Up to total derivatives, the most general such action with at most two time derivatives may be parametrized in terms of a mass parameter m^2 and an overall dimensionless normalization λ_ϕ :

$$I_{(\phi^2)} = \frac{1}{2\ell_P^{d-1}} \int \sqrt{-\bar{g}} \lambda_\phi \left\{ -\frac{1}{2} ((\bar{\nabla}\phi)^2 + m^2\phi^2) \right\}. \quad (2.26)$$

This action defines the background dynamics of ϕ as $(\bar{\square} - m^2)\phi = 0$.

- **Backreaction on the geometry:** there are several possible couplings of the type $h\phi\phi$, which are inequivalent upon integration by parts. We can parametrize them as follows:

$$I_{(h\phi^2)} = \frac{1}{2\ell_P^{d-1}} \int \sqrt{-\bar{g}} \left\{ -\frac{h}{4} (c_1 (\bar{\nabla}\phi)^2 + c_2 m^2 \phi^2) + \frac{c_3}{2} h^{ab} \bar{\nabla}_a \phi \bar{\nabla}_b \phi \right. \\ \left. + c_4 h^{ab} \phi \bar{\nabla}_a \bar{\nabla}_b \phi + c_5 h \phi \bar{\square} \phi \right\}, \quad (2.27)$$

Clearly the coefficients c_2 and c_5 are on-shell equivalent. We will thus set $c_5 = 0$ in what follows. Now demanding diffeomorphism invariance, an argument completely analogous to the one we gave for (2.23) implies the following conditions for the remaining parameters:¹¹

$$c_1 = c_2 = c_3, \quad c_4 = 0. \quad (2.28)$$

which leaves us with only one free parameter c_1 . However, it is immediately clear that the overall normalization has to be the same as in (2.26), i.e.,

$$c_1 = \lambda_\phi. \quad (2.29)$$

If this normalization was any different, h_{ab} would not be the response to a scalar with mass m^2 and normalization λ_ϕ . Taking into account these constraints, we can rewrite (2.27) as

$$I_{(h\phi^2)} = \frac{1}{2\ell_P^{d-1}} \int \sqrt{-\bar{g}} \frac{\lambda_\phi}{2} h^{ab} T_{ab}^{(\phi)}, \quad (2.30)$$

with $T_{ab}^{(\phi)} \equiv -\frac{1}{2} [(\bar{\nabla}\phi)^2 + m^2\phi^2] \bar{g}_{ab} + \bar{\nabla}_a \phi \bar{\nabla}_b \phi,$

¹¹Alternatively, one can determine the coefficients by demanding conservation of the matter stress tensor, $\bar{\nabla}^b T_{ab}^{(\phi)} = 0$.

Summarizing these two points, we recognize the most general scalar-graviton couplings which are consistent with the perturbative scheme and diffeomorphism invariance as being exactly the same as the expansion of the following action up to $\mathcal{O}(h)$:

$$\begin{aligned} I_{\text{matter}} &= I_{(\phi^2)} + I_{(h\phi^2)} + \mathcal{O}(h^2) \\ &= \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \lambda_{\phi} \left[-\frac{1}{2} ((\nabla\phi)^2 + m^2\phi^2) \right] + \mathcal{O}(h^2), \end{aligned} \quad (2.31)$$

where $g_{ab} = \bar{g}_{ab} + h_{ab}$. Of course, this result is not too surprising as this is the standard action for a massive scalar minimally coupled to gravity.

Most general second order action. We can summarize the above discussion as follows. The most general way of coupling h_{ab} to a scalar ϕ in a way that is consistent with our perturbative scheme and respects diffeomorphism invariance and some very basic constraints coming from the particular CFT computation we are interested in, is given by (2.21). This action has only four dimensionless parameters $\{\ell_{\text{AdS}}/\ell_{\text{P}}, m^2\ell_{\text{AdS}}^2, \lambda_{\phi}, \lambda_h\}$. We already explained in section 2.3.1 how for our problem the first three of these have to be determined holographically in terms of the dimensionless CFT quantities $\{a_d^*, \Delta, C_{\mathcal{O}}\}$. We will argue shortly that the fourth parameter λ_h does not enter the computation of interest and therefore its value is irrelevant for computing perturbed entanglement entropy. In addition, there is, of course, the boundary coupling λ which enters the gravitational setup by prescribing the boundary condition for ϕ .

As a consequence, any covariant theory of gravity whose linearization takes the form (2.21), provides a good auxiliary system for computing $S_{EE}(\mathcal{A})$ at second order in perturbation theory. For instance, the action (2.21) with $\lambda_h = 1$ is precisely the same as the linearization of the cosmological Einstein-scalar action (2.17) via $g_{ab} = \bar{g}_{ab} + h_{ab}$. There are, of course, many other non-linear actions whose linearization on maximally symmetric backgrounds also takes the form (2.21) with a different value of λ_h , e.g., those of Lovelock type [46] or quasi-topological gravity [47]. Indeed, even if the CFT in question is actually dual to some particular higher derivative theory; for the purpose of calculating the perturbed entanglement entropy, one can always work with the Einstein action (2.17) after suitably matching the parameters $\{\ell_{\text{AdS}}/\ell_{\text{P}}, m^2\ell_{\text{AdS}}^2, \lambda_{\phi}\}$ with the CFT boundary conditions. The different linearized actions will then only differ by their value of λ_h . We will see examples of this in section 3.

Since the normalization λ_h is the only ingredient in (2.21) which would genuinely distinguish bulk theories with higher derivative couplings from those without, we now need to understand why the linearized computation of perturbed entanglement entropy of a ball is not sensitive to this.

Holographic interpretation of λ_h . Before we continue with the gravitational analysis, it is instructive to elaborate on the meaning of λ_h , which has a very natural interpretation from the CFT point of view. The graviton normalization λ_h corresponds to the normalization of the CFT stress tensor two-point function, usually denoted by C_T . It sets the

strength of the leading divergence in the two-point function of CFT stress tensors:

$$\langle T_{\mu\nu}(x)T_{\rho\sigma}(0) \rangle = \frac{C_T}{x^{2d}} \mathcal{I}_{\mu\nu,\rho\sigma}(x), \quad (2.32)$$

with some well known universal function $\mathcal{I}_{\mu\nu,\rho\sigma}(x)$ [33, 34]. In $d = 2$, the normalization agrees with the usual c -central charge, $C_T = c$. Similarly, in $d = 4$ one finds $C_T = (40/\pi^2)c$. Once a holographic description of the CFT has been given, usual AdS/CFT methods can be used to compute the left hand side of (2.32) as a bulk graviton correlator. Since λ_h sets the normalization of the graviton propagator in the presence of matter sources (cf., (2.38)), it is rather clear that C_T will be proportional to λ_h . Doing the holographic calculation in detail leads to the standard result [48]

$$C_T = \frac{f_d}{2} \frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \lambda_h, \quad \text{where} \quad f_d \equiv \frac{2d(d+1)}{d-1} \frac{\Gamma(d)}{\pi^{d/2}\Gamma(d/2)}. \quad (2.33)$$

It is now obvious that once $\ell_{\text{AdS}}/\ell_{\text{P}}$ has been fixed such as to reproduce the other central charge a_d^* via matching (2.7), there is no free parameter left in (2.33) which could absorb the normalization λ_h . In our setup, gravity theories with different λ_h will hence describe CFTs with different C_T . Conversely, a CFT measurement of C_T would allow to for a distinction between different bulk models. However, we will now argue that the perturbed entanglement entropy of spheres is agnostic about this.

2.3.3 Second order Wald entropy is independent of graviton normalization

As we have seen, any gravity theory whose action linearized around AdS_{d+1} takes the form (2.21) can be used to compute second order perturbations of entanglement entropy of spheres. However, we identified one parameter λ_h in the action (2.21) which sets the graviton normalization and turns out to be sensitive to more detailed bulk structures such as higher curvature terms. We will now argue that this parameter does not enter the second order entanglement entropy.¹² This is the gravity version of the statement that (2.16) is independent of C_T .

We recall that due to the spherical symmetry of \mathcal{A} , extrinsic curvatures on the minimal surface \mathcal{E}_{min} vanish and the entanglement entropy functional to be evaluated on \mathcal{E}_{min} is just the Wald functional. In vacuum this reduces to just the area of \mathcal{E}_{min} multiplied by a normalization which is basically determined by a_d^* as in (2.5). The perturbation can then be computed in holography as

$$\delta S_{EE}(\mathcal{A}) = \delta S_{\text{Wald}}(\mathcal{E}_{\text{min}}) = \delta \left(-2\pi \int_{\mathcal{E}_{\text{min}}} \sqrt{\gamma} \frac{\delta \mathcal{L}}{\delta R_{abcd}} n_{ab} n_{cd} \right). \quad (2.34)$$

Clearly $\delta S_{\text{Wald}}(\mathcal{E}_{\text{min}})$ can be expanded perturbatively by considering different powers of h_{ab} and ϕ . From the structure of the gravitational backreaction it follows that there is no contribution at $\mathcal{O}(\lambda)$. At $\mathcal{O}(\lambda^2)$ (i.e., at first order in h_{ab}) the variation in (2.34) acts on the

¹²See section C for related observations based on the entanglement first law.

objects inside the integral, but it does not change the surface \mathcal{E}_{\min} itself.¹³ Since the surface \mathcal{E}_{\min} is maximally symmetric, we anticipate that this variation of the Wald functional is just a variation of the area form (up to normalization). That is, the evaluation of (2.34) should yield precisely the same answer as in Einstein gravity up to normalization. Let us make this slightly more explicit to figure out the correct normalization.

To evaluate (2.34) at $\mathcal{O}(\lambda^2)$, we should expand the integrand linearly in the perturbation:

$$\delta S_{\text{Wald}}(\mathcal{E}_{\min}) = -2\pi \int_{\mathcal{E}_{\min}} \left[\sqrt{\gamma} \frac{\delta \mathcal{L}}{\delta R_{abcd}} n_{ab} n_{cd} \right]_{\mathcal{O}(h)}. \quad (2.35)$$

The obvious way to evaluate this is to start with the full non-linear Lagrangian, expand the Wald entropy density around the AdS background, and then identify the $\mathcal{O}(h)$ term in (2.35). However, even without knowledge of the full theory it is clear that ultimately the integrand can only be sensitive to the part of \mathcal{L} which is at most quadratic in h_{ab} , i.e., we should be able to predict the answer just based on the most general second order Lagrangian (2.21). Moreover, only the purely gravitational sector of (2.21) is relevant for Wald entropy at $\mathcal{O}(\lambda^2)$. This is quite clear from the observations that a minimally coupled scalar field in gravity does not couple to curvature and the action (2.21) agrees with that of a minimally coupled scalar at the order we are interested in.

From these observations, we are led to conclude that the integrand in (2.35) formally has the same structure in any allowed theory of gravity up to the normalization λ_h . Therefore, without loss of generality, we can compute the integral as if it came from Einstein's theory and multiply by λ_h . But in that case, the answer for $\delta S_{\text{Wald}}(\mathcal{E}_{\min})$ is well known: it is just the variation of the area functional. Therefore, (2.35) can be rewritten in general as a variation of area, multiplied by the normalization λ_h that encodes the higher derivative dependence:

$$\delta S_{\text{Wald}}(\mathcal{E}_{\min}) = 2\pi \left(\frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \lambda_h \right) \frac{\int_{\mathcal{E}_{\min}} \sqrt{\gamma} \left(\frac{1}{2} \bar{\gamma}^{\alpha\beta} h_{\alpha\beta} \right)}{\ell_{\text{AdS}}^{d-1}} + \mathcal{O}(h^2). \quad (2.36)$$

where $\bar{\gamma}_{\alpha\beta}$ is the background metric induced on \mathcal{E}_{\min} by \bar{g}_{ab} and $h_{\alpha\beta}$ is the induced metric perturbation.¹⁴ Note that the arguments just given were heuristic and we are discarding boundary terms which would be important to regulate the divergences. A detailed derivation of (2.36) can be found in section B.

The integral (2.36) can now be used to derive the independence of $\delta S_{\text{Wald}}(\mathcal{E}_{\min})$ on λ_h . We just need to find the on-shell dependence of $h_{\alpha\beta}$ on λ_h and combine it with the

¹³Variations of the bifurcation surface only affect the Wald entropy at the next order, which is beyond our discussion. To see this, note that the surface \mathcal{E}_{\min} extremizes the Wald entropy functional upon variation with respect to the metric. First order variations of the metric (which are $\mathcal{O}(\lambda^2)$ in our conventions) hence leave \mathcal{E}_{\min} unchanged.

¹⁴Explicitly, if $\xi^\alpha(x)$ denotes the coordinates of \mathcal{E}_{\min} embedded in AdS_{d+1} , we have

$$\bar{\gamma}_{\alpha\beta} = \frac{\partial x^a}{\partial \xi^\alpha} \frac{\partial x^b}{\partial \xi^\beta} \bar{g}_{ab}, \quad h_{\alpha\beta} = \frac{\partial x^a}{\partial \xi^\alpha} \frac{\partial x^b}{\partial \xi^\beta} h_{ab}. \quad (2.37)$$

For more conventions see section B.

prefactor in (2.36). To this end, consider the linearized equations of motion implied by the general action (2.21):

$$\lambda_h \mathcal{G}_{ab}{}^{cd} h_{cd} = -\frac{\lambda_\phi}{2} T_{ab}^{(\phi)}. \quad (2.38)$$

From this equation it is clear how to go from the universality of graviton propagation to the universality of perturbed Wald entropy: apart from the overall λ_h every dimensionless parameter in (2.38) has been fixed by the boundary input $\{a_d^*, \Delta, C_{\mathcal{O}}\}$. The combination $\lambda_h h_{cd}$ therefore satisfies a universal equation of motion which is independent of the non-linear details of the bulk theory. Since the perturbed Wald entropy (2.36) depends on precisely this particular combination it is clear that $\delta S_{\text{Wald}}(\mathcal{E}_{\text{min}})$ will be the same in all theories we consider. Said differently, the function h_{ab} which satisfies the backreaction equation (2.38) is proportional to $1/\lambda_h$ such that the λ_h dependence in (2.36) cancels. We conclude that $\delta S_{EE}(\mathcal{A}) = \delta S_{\text{Wald}}(\mathcal{E}_{\text{min}})$ is independent of λ_h . This is consistent with the field theory result (2.16) being independent of C_T .

This concludes our analysis of the role which various dimensionless parameters play. Having verified the cancellation of λ_h , the explicit solution of (2.38) and evaluation of (2.36) can be done without loss of generality at $\lambda_h = 1$, i.e., Einstein gravity. We refrain from repeating this calculation. See, for example, [14, 15] for an explicit evaluation of (2.36) in the case of a scalar field source. The result of their computation perfectly matches the CFT answer (2.16).

3 Example: Gauss-Bonnet theory

In this section we wish to illustrate the general argument, using the example of gravity theories with Lagrangians quadratic in curvature. This discussion will make contact with a number of previous higher derivative computations of entanglement entropy. In particular we will show that various results for perturbations of entanglement entropy of spherical regions dual to higher derivative theories can be equally well reproduced from a calculation in cosmological Einstein gravity by renormalizing the discrete set of free parameters which play a role at second order in perturbation theory, i.e., $\{\ell_{\text{AdS}}/\ell_{\text{P}}, m^2 \ell_{\text{AdS}}^2, \lambda_\phi\}$. We will also see the appearance of the additional parameter λ_h which drops out of the formula for entropy.

3.1 Two gravity models

The basic gravitational model that we want to take as the reference system is just the auxiliary Einstein-Hilbert theory described by

$$I = \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \left[R + \frac{d(d-1)}{\ell^2} - \frac{\lambda_\phi}{2} ((\nabla\phi)^2 + m^2 \phi^2) \right], \quad (3.1)$$

with asymptotic boundary condition $\phi(z \rightarrow 0) \simeq \lambda z^{d-\Delta}$ such that ϕ sources a scalar operator with dimension Δ and coupling λ in the CFT. We parametrize the lowest order backreaction of the scalar probe on the geometry as $g_{ab} = \bar{g}_{ab} + h_{ab}$, where \bar{g}_{ab} is AdS_{d+1} with scale $\ell_{\text{AdS}}^2 = \ell^2$. The independent dimensionless parameters in the gravitational problem can therefore be taken as $\{(\ell_{\text{AdS}}/\ell_{\text{P}}), m^2 \ell_{\text{AdS}}^2, \lambda_\phi\}$, while $\lambda_h = 1$ for Einstein gravity.

In this section we wish to reproduce predictions of four-derivative theories by using the above Einstein theory with renormalized parameters. The only four-derivative theory which passes the consistency condition of yielding second order equations of motion is Gauss-Bonnet theory, which we will discuss here. Nevertheless, in section A.1 we explore the most general four-derivative theory of gravity and come to very similar conclusions in the case where the higher derivative terms are small perturbative corrections to (3.1).

With this motivation in mind, let us take the following action as the theory to compare with:

$$\tilde{I} = \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \left[R + \frac{d(d-1)}{\tilde{\ell}^2} + \tilde{\ell}^2 \mathcal{L}_{h.d.} - \frac{\tilde{\lambda}_\phi}{2} ((\nabla\phi)^2 + \tilde{m}^2 \phi^2) \right] \quad (3.2)$$

where $\mathcal{L}_{h.d.} = \gamma (R^2 - 4R_{ab}R^{ab} + R_{abcd}R^{abcd})$

again with a fixed boundary condition $\phi(z \rightarrow 0) \simeq \lambda z^{d-\Delta}$. The AdS_{d+1} background solution of this theory will have some other scale $\tilde{\ell}_{\text{AdS}}$ which is a function of $\tilde{\ell}$ and γ . In the context of both theories I and \tilde{I} , we will use \bar{g}_{ab} to refer to the AdS_{d+1} background. From the context it should be clear whether \bar{g}_{ab} is the background with ℓ_{AdS} or with $\tilde{\ell}_{\text{AdS}}$.

Our basic aim is to show that for the purpose of carrying out the perturbed entanglement entropy computation of spherical regions, the actions (3.1) and (3.2) are indistinguishable if the dimensionless parameters $\{(\ell_{\text{AdS}}/\ell_{\text{P}}), m^2 \ell_{\text{AdS}}^2, \lambda_\phi\}$ are suitably expressed as functions of $\{(\tilde{\ell}_{\text{AdS}}/\ell_{\text{P}}), \tilde{m}^2 \tilde{\ell}_{\text{AdS}}^2, \tilde{\lambda}_\phi, \gamma\}$ such that they are consistent with the CFT boundary conditions. In order to show this, we will now derive the relevant ingredients for such a computation order by order in λ from both I and \tilde{I} . We will demonstrate that the parameters of Einstein gravity can always be chosen such that the two perturbative expansions match. The higher derivative coupling γ can be absorbed entirely into renormalizations of the other dimensionless quantities.

3.2 Parameter matching

We will now determine the relation between the dimensionless parameters of the models I and \tilde{I} such that both make identical predictions for vacuum and perturbed entanglement entropy of spheres.

Finding the AdS_{d+1} background solution. Before turning on the scalar perturbation, we need to determine the maximally symmetric AdS_{d+1} background solutions of the two theories (3.1) and (3.2). It is well known that such solutions exist in both theories and are ghost-free [49] if we choose the respective AdS scales

$$\begin{aligned} \ell_{\text{AdS}}^2 &= \ell^2, \\ \tilde{\ell}_{\text{AdS}}^2 &= \frac{\tilde{\ell}^2}{2} \left(1 + \sqrt{1 - 4(d-2)(d-3)\gamma} \right). \end{aligned} \quad (3.3)$$

We will henceforth expand around the AdS_{d+1} background and thus treat ℓ_{AdS} and ℓ as equivalent (and similarly with tildes).

Matching of Wald entropy (or a_d^*). Let us remind the reader of the result of section 1 that Wald entropy of maximally symmetric spaces is always proportional to area. We can thus compute the unperturbed entanglement entropy of \mathcal{A} as the Wald entropy of hyperbolic black holes corresponding to the two AdS radii (3.3) of Einstein and Gauss-Bonnet gravity, respectively. By integrating (2.5) for the two gravity models, one finds

$$\begin{aligned} S_{\text{Wald}} &= \frac{2\pi}{\pi^{d/2}} \Gamma(d/2) a_d^* \frac{\text{Vol}(\mathbb{H}_{d-1})}{\ell_{\text{AdS}}^{d-1}}, \quad a_d^* = \frac{\pi^{d/2}}{\Gamma(d/2)} \frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}}, \\ \tilde{S}_{\text{Wald}} &= \frac{2\pi}{\pi^{d/2}} \Gamma(d/2) \tilde{a}_d^* \frac{\text{Vol}(\mathbb{H}_{d-1})}{\tilde{\ell}_{\text{AdS}}^{d-1}}, \quad \tilde{a}_d^* = \frac{\pi^{d/2}}{\Gamma(d/2)} \frac{\tilde{\ell}_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \left(1 - 2(d-1)(d-2) \gamma \frac{\tilde{\ell}^2}{\tilde{\ell}_{\text{AdS}}^2} \right). \end{aligned} \quad (3.4)$$

Given the entanglement entropy $S_{EE}^{(0)}(\mathcal{A})$ with fixed normalization in the CFT, it can be reproduced by the computation of Wald entropy in either of the gravity models by requiring $a_d^* = \tilde{a}_d^*$, i.e.,

$$\frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} = \frac{\tilde{\ell}_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \left(1 - 2(d-1)(d-2) \gamma \frac{\tilde{\ell}^2}{\tilde{\ell}_{\text{AdS}}^2} \right). \quad (3.5)$$

This determines the dimensionless ratio $\ell_{\text{AdS}}/\ell_{\text{P}}$ in Einstein gravity in terms of $\{\tilde{\ell}_{\text{AdS}}/\ell_{\text{P}}, \gamma\}$ in Gauss-Bonnet theory such that both theories make the same predictions for Wald entropy of the maximally symmetric horizon. From the CFT point of view, this matching is the condition that a_d^* is a fixed fundamental parameter of the CFT which should be reproduced by either candidate of a gravity dual.

Matching of scalar coupling (or $C_{\mathcal{O}}$). Perturbing the CFT with \mathcal{O} means introducing a massive scalar in the gravitational action. The mass of the scalar is constrained by the requirement that the boundary value of the scalar sources the operator \mathcal{O} with given dimension Δ . This fixes

$$\ell_{\text{AdS}}^2 m^2 = \Delta(\Delta - d) = \tilde{\ell}_{\text{AdS}}^2 \tilde{m}^2. \quad (3.6)$$

The dynamics of the scalar field is derived at lowest order from the propagation on AdS_{d+1} vacuum:

$$(\bar{\square} - m^2) \phi = 0, \quad \phi(z \rightarrow 0) \sim \lambda z^{d-\Delta}, \quad (3.7)$$

and similarly with tildes. Once the boundary condition is imposed in this way, standard holographic techniques lead to the matching between the normalization $C_{\mathcal{O}}$ of the relevant CFT correlators and the normalization of the scalar field action via (2.20): the two gravity theories I and \tilde{I} lead to the same $C_{\mathcal{O}}$ provided that

$$\lambda_{\phi} \frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} = \tilde{\lambda}_{\phi} \frac{\tilde{\ell}_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}}, \quad (3.8)$$

so the ratio $\lambda_{\phi}/\tilde{\lambda}_{\phi}$ should be fixed to be the same as in (3.5).

Computation of perturbed entanglement entropy. We finally want to check explicitly the arguments of section 2.3.3 and verify that after the matching conditions above have been implemented, second order entanglement entropy is automatically the same for

I and \tilde{I} . We will also verify that the difference between $\lambda_h = 1$ and $\tilde{\lambda}_h$ cannot be absorbed in any further free parameters. This will lead to the conclusion that the central charge C_T is genuinely different in the CFTs dual to I compared to those dual to \tilde{I} after the above matchings have been performed.

Our perturbative scheme in the CFT demands that we consider the linearized equations of motion in the bulk. The linearized equations of motion can be obtained from the actions (3.1), (3.2) by linearizing around $\bar{g}_{\mu\nu} = g_{\mu\nu} - h_{\mu\nu}$. We find for I and \tilde{I} , respectively [50]:

$$\begin{aligned}\lambda_h \mathcal{G}_{ab}{}^{cd} h_{cd} &= -\frac{\lambda_\phi}{2} T_{ab}^{(\phi)}, \\ \tilde{\lambda}_h \tilde{\mathcal{G}}_{ab}{}^{cd} h_{cd} &= -\frac{\tilde{\lambda}_\phi}{2} \tilde{T}_{ab}^{(\phi)},\end{aligned}\tag{3.9}$$

where $T_{ab}^{(\phi)}$ and $\tilde{T}_{ab}^{(\phi)}$ are the source terms (2.30) for scalar matter with mass m^2 and \tilde{m}^2 , respectively propagating on AdS_{d+1} with scale ℓ_{AdS} and $\tilde{\ell}_{\text{AdS}}$. The kinetic operators \mathcal{G}_{abcd} and $\tilde{\mathcal{G}}_{abcd}$ are defined as in (2.22), but, of course, also refer to backgrounds with respective scales ℓ_{AdS} and $\tilde{\ell}_{\text{AdS}}$. The coefficients λ_h and $\tilde{\lambda}_h$ in (3.9) can be computed for I and \tilde{I} theories and turn out to be

$$\lambda_h = 1, \quad \tilde{\lambda}_h = \left(1 - 2(d-2)(d-3) \gamma \frac{\tilde{\ell}^2}{\ell_{\text{AdS}}^2}\right).\tag{3.10}$$

Clearly there is no free parameter left in (3.9) which could be renormalized to absorb these normalizations. From (2.33) it is hence clear that the C_T charges of the dual CFTs will differ by the ratio $\lambda_h/\tilde{\lambda}_h$. At this stage, any CFT observable which did depend on C_T would therefore genuinely distinguish between the two gravity models. As shown before, $\delta S_{EE}(\mathcal{A})$ does not depend on C_T , though.

Our matching prescriptions (3.3), (3.5) and the graviton kinetic term scaling (3.10) have, of course, appeared in the literature before. For example, see [20] for a computation of perturbed sphere entanglement entropy in Gauss-Bonnet theory. One can easily verify that their computation gives an answer which takes the form of perturbed entanglement entropy in Einstein gravity with dimensionless parameters rescaled as predicted by our general analysis.¹⁵ At this order the entanglement entropy computation for spherical regions is therefore not sensitive to the Gauss-Bonnet coupling.

4 Discussion

Let us summarize our arguments and comment on implications for the program of bulk reconstruction.

¹⁵See, e.g., eq. (7.7) of [20] with the understanding that their Gauss-Bonnet coupling λ equals 2γ in our notation.

Summary. We considered the second order perturbation of CFT vacuum entanglement entropy of spheres due to scalar operator deformations. This quantity exhibits a universal functional form which only depends on a small number of fundamental CFT parameters to set various normalizations. We explored a similar universality statement about gravity, which explains why the corresponding holographic computation can always be recast in terms of one universal gravitational system, viz., linearized Einstein gravity coupled to a free massive scalar field. We argued that any ghost-free gravitational theory other than Einstein gravity can be used to make identical predictions due to the uniqueness of diffeomorphism invariant spin-2 propagation on AdS. For instance, we demonstrated how in Gauss-Bonnet theory both the vacuum and the perturbed sphere entanglement entropy (obtained by computing Wald entropy of a minimal surface) can be reproduced from just Einstein gravity by suitably renormalizing the dimensionless parameters of the theory. In this sense, knowledge of this entanglement data does not allow to distinguish between Gauss-Bonnet theory and Einstein gravity.

More precisely, the perturbed sphere entanglement only depends on the dimension Δ of the deforming operator \mathcal{O} , the generalized a -central charge a_d^* (in vacuum) and the normalization $C_{\mathcal{O}}$ of the $\langle \mathcal{O}\mathcal{O} \rangle$ correlator (at second order in the perturbation). We saw that any viable gravity model has enough dimensionless parameters to account for these dependencies (the scalar mass $\ell_{\text{AdS}}^2 m^2$, the normalization of the gravitational action, $\ell_{\text{AdS}}/\ell_{\text{P}}$, and the normalization of the scalar field, λ_{ϕ} , respectively). In particular, no higher curvature terms in the bulk are required in order to obtain a sufficient number of couplings for consistently performing the holographic matching. In this problem, the only signature of higher curvature couplings would be a renormalization of the parameters of cosmological Einstein gravity. This implies that perturbed entanglement entropy of spheres can be computed using any ghost-free gravitational model as an auxiliary system, even for CFTs without any simple gravity dual at all [11].

In face of this, it is clear that for the derivation of bulk dynamics from entanglement in the sense of AdS/CFT it is desirable to reconstruct bulk features which are not entirely universal in the above sense. The obvious question arises which quantities one should consider instead of or in addition to entanglement entropy of spheres in order to achieve a dependence on bulk features which are not the same for all CFTs. We will come back to this question below.

Generalizations. There are some modifications of our analysis which would be interesting to explore. For instance, one could compute the perturbation to higher order in λ . By extending the perturbation theory of section 2.2, it is clear that entanglement entropy will then be sensitive to higher point correlation functions which would probe bulk dynamics beyond the linear approximation. While this would be very interesting to see explicitly, the major complication in the bulk would be a backreaction not just on the Wald entropy density, but on the minimal surface itself. Once the minimal surface has non-vanishing extrinsic curvatures, ambiguities in the entanglement entropy functional need to be addressed [21–24]. By resolving this issue, one may hope to derive bulk dynamics beyond linear order.

We mainly focused on perturbations of entanglement due to deformations of the CFT by a relevant scalar operator. As another generalization, one can imagine deformations by operators of other spin. As long as the deformation does not involve a stress tensor, our conclusions are expected to hold since the new operator comes at least with an overall normalization which is all that is required to account for the renormalization of $\ell_{\text{AdS}}/\ell_{\text{P}}$. Consider, however, the case where the CFT is deformed by a stress tensor insertion. In that case, the first interesting perturbation of the entanglement entropy occurs at $\mathcal{O}(\lambda)$ instead of $\mathcal{O}(\lambda^2)$, where one gets a contribution $\sim \langle TT \rangle$ with normalization set by C_T . In the bulk the problem of gravitational backreaction changes in a similar way and the extremal surface will feel a perturbation at $\mathcal{O}(\lambda)$. It would be interesting to explore this case in detail and consider the dependence on C_T to draw similar conclusions about the effect of higher curvature couplings. The dependence on C_T is expected to lead to a way of distinguishing Einstein gravity duals from higher curvature ones. In fact it has been argued before (as we review in section C) that generic perturbations of spherical region entanglement entropy can be used to derive linearized bulk equations including their normalization set by C_T [6].

Further generalizations could involve perturbations of the CFT other than operator deformations. For instance, a small geometric perturbation of the spatial sphere should lead to similar conclusions as those just mentioned. Similarly, $1/N$ corrections have recently been explored in order to improve on the entanglement first law [51]. It would be interesting to pursue this further and explore ways to reconstruct bulk dynamics at a level where one can distinguish between different higher curvature theories. Instead of considering individual approaches in detail, let us now outline in general how a good choice of CFT observables can be made which are sensitive to interesting bulk features in a controlled way.

How to “bootstrap” bulk theories. As we have just seen, there are, of course, simple ways to circumvent the problem of perturbed sphere entanglement being insensitive to the desired detailed bulk information. Roughly speaking, one needs to consider more detailed CFT structures in order to distinguish between different bulk models. In general one could imagine bootstrapping the bulk theory using the CFT input. Clearly it is hard to derive bulk dynamics directly from fundamental CFT parameters such as two-point functions and OPE coefficients. But entanglement entropy seems to provide a physically meaningful quantity with well-understood universal properties on both sides of the duality, which depends on these fundamental building blocks of the CFT in a useful way. At lowest order, the vacuum entanglement entropy of spheres can be used to measure the generalized a -charge a_d^* . In our particular context we used this to fix the normalization $\ell_{\text{AdS}}/\ell_{\text{P}}$ of any possible bulk dual (whose dynamics is uniquely fixed in the same way as the functional form of entanglement entropy is universal). As we have seen, a small perturbation of this entanglement entropy does not affect the dynamics enough to determine if there are any higher curvature terms present in the bulk theory. To answer this question, one would have to measure some other entanglement entropy which is sensitive to further details of the bulk dynamics.

One natural CFT parameter that we have only encountered briefly so far is the stress tensor central charge. Let us elaborate how any quantity which is sensitive to C_T will also be sensitive to bulk higher derivative couplings and hence will not be reproducible by just

Einstein gravity once $\ell_{\text{AdS}}/\ell_{\text{P}}$ has been fixed by matching vacuum entanglement entropy of spheres as outlined in this note. In our context, the ratio a_d^*/C_T computed holographically is independent of $\ell_{\text{AdS}}/\ell_{\text{P}}$ and genuinely takes different values in the presence of higher curvature terms. In fact it is clear from (2.7) and (2.33) that higher curvature terms in the bulk are necessary whenever $a_d^*/C_T = 2\pi^{d/2}/\Gamma(d/2)f_d$ does *not* hold. This is, of course, just a generalization of the well-known statement that higher curvature terms in the bulk theory are required whenever the a - and c -central charges parametrizing the boundary conformal anomaly differ [52–54].

Any set of CFT observables which depends on both a_d^* and C_T can hence be used to detect if higher curvature terms in the bulk are necessary. The problem of genuine bulk reconstruction is therefore not at all the restriction to quantities which are universal in the sense that they only depend on an overall normalization. But it is now clear that one has to consider *enough* such quantities in order to see a necessity to account for their values in terms of higher curvature interactions.

There are many examples of entanglement-related CFT quantities which one may envisage to compute in this context. For instance, the entanglement entropies of other surfaces such as infinite strips or cylinders are sensitive to both a_d^* and C_T [19, 20, 55]. Further, the presence of corners and kinks in the spatial region of interest leads to divergences which exhibit a universal dependence on the opening angle and (at least in holographic CFTs) depend on a_d^* and C_T in an interesting way [56–63]. Other interesting quantities with universal features and dependence on C_T are Rényi entropies of spherical regions (e.g., in the entanglement entropy limit $q \rightarrow 1$) [64–66]. All these quantities have in common a contribution with a universal functional form and a dependence on few CFT parameters (in particular a_d^* and C_T). They could hence be used to distinguish between CFTs with a dual description in terms of Einstein gravity and CFTs which require higher curvature terms in the bulk.

However, note that even if some CFT measurement of a_d^* and C_T leads to the conclusion that there can exist a semiclassical bulk theory only if contains higher curvature interactions are present, it is usually not possible to determine precisely what higher curvature terms are required. Indeed, the signature of, say, a Gauss-Bonnet term on a_d^*/C_T can equally well be reproduced by some other higher derivative term. For illustration of this statement see also section B, where it is shown how any higher curvature theory simplifies to the background field expansion of Gauss-Bonnet form if one is only concerned with linearized equations of motion. In this sense Gauss-Bonnet theory is the most general theory which can account for arbitrary values of a_d^* and C_T in the CFT. In order to further bootstrap the precise bulk dynamics, one needs again access to more CFT data (e.g., entanglement data which depends on further 2-point functions or OPE coefficients). The reason is, again, that gravitons propagating on AdS have very universal dynamics. This is a manifestation of the familiar problem that entanglement entropies which are computable tend to be very universal and contain only discrete information about the actual bulk dynamics.

It would be interesting to use this philosophy to derive more detailed (non-linear) bulk dynamics for some simple CFT with known holographic dual from first principles. Generalizing this approach may then ultimately lead to a genuine reconstruction of the bulk geometry and its dynamics by using a suitable set of entanglement entropies.

Acknowledgments

I am very grateful to Jyotirmoy Bhattacharya and Mukund Rangamani for collaboration during early stages of this project, numerous insightful discussions and comments on a draft. I also wish to thank Rob Myers for useful comments. I am supported by a Durham Doctoral Fellowship.

A Generalization to bulk theories with higher order equations of motion

In the main text, we considered the case where the equations of motion for the metric perturbation in the bulk were strictly second order. The motivation for doing so was to exclude ghost modes which would occur if the graviton equations were higher than second order. However, if the higher curvature contributions to gravity are parametrically small, they can be interpreted, e.g., as toy models for perturbative string theory corrections to Einstein gravity. The ghost modes are then an artefact of the low energy truncation and can be ignored, which allows for a considerable increases of the number of acceptable gravitational models covered by our analysis. In order to illustrate this, we first generalize in section A.1 the Gauss-Bonnet theory of section 3 to the case of all Lagrangians which are quadratic in curvature. In section A.2 we then give a more abstract argument to show that this example actually illustrates the general case.

A.1 Example: general curvature squared theories

In this section we generalize the analysis of section 3 to theories where the gravitational sector is described by

$$\tilde{I} = \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \left[R + \frac{d(d-1)}{\tilde{\ell}^2} + \tilde{\ell}^2 \mathcal{L}_{h.d.} - \frac{\tilde{\lambda}_\phi}{2} ((\nabla\phi)^2 + \tilde{m}^2 \phi^2) \right] \quad (\text{A.1})$$

where $\mathcal{L}_{h.d.} = \alpha R^2 + \beta R_{ab}R^{ab} + \gamma (R^2 - 4R_{ab}R^{ab} + R_{abcd}R^{abcd})$

We do this to explicitly demonstrate that such theories — despite the presence of massive ghost modes — are acceptable in a perturbative sense. Indeed, we will see how the ghost degrees of freedom are suppressed and hence can be neglected for the computations we want to perform.

We want to model the theory (A.1) with the cosmological Einstein-scalar theory (3.1). That is, we will give the matching conditions which express the parameters of the latter theory, $\{(\ell_{\text{AdS}}/\ell_{\text{P}}), m^2 \ell_{\text{AdS}}^2, \lambda_\phi\}$, in terms of $\{(\tilde{\ell}_{\text{AdS}}/\ell_{\text{P}}), \tilde{m}^2 \tilde{\ell}_{\text{AdS}}^2, \tilde{\lambda}_h, \tilde{\lambda}_\phi, \alpha, \beta, \gamma\}$ such that the theories make the same predictions for entanglement entropy of spheres in vacuum even after deforming the theory with a scalar operator. The main new ingredient is the presence of four-derivative terms in the equations of motion.

Before discussing the linearized equations of motion, let us give the analogue of the matching conditions (3.3), (3.5). The stable AdS_{d+1} background solutions in the two

theories have the following scales in terms of the cosmological constant:

$$\begin{aligned}\ell_{\text{AdS}}^2 &= \ell^2, \\ \tilde{\ell}_{\text{AdS}}^2 &= \frac{\tilde{\ell}^2}{2} \left(1 + \sqrt{1 - \frac{4(d-3)}{(d-1)} (d(d+1)\alpha + d\beta + (d-1)(d-2)\gamma)} \right).\end{aligned}\quad (\text{A.2})$$

The generalization of the a_d^* matching condition (3.5) then becomes

$$\frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} = \frac{\tilde{\ell}_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \left(1 - 2(d(d+1)\alpha + d\beta + (d-1)(d-2)\gamma) \frac{\tilde{\ell}^2}{\tilde{\ell}_{\text{AdS}}^2} \right). \quad (\text{A.3})$$

Exactly as in section 3.2, the previous equation also immediately fixes the ratio of scalar couplings, i.e., $\tilde{\lambda}_\phi/\lambda_\phi$ is given by the value bracket on the right hand side of (A.3). This guarantees that holographic computations of $\langle \mathcal{OO} \rangle$ yield a unique normalization $C_{\mathcal{O}}$.

Consider now the linearized graviton equations in the two theories. For this, we find it convenient to gauge fix the metric. In the transverse traceless gauge ($\bar{\nabla}_a h^{ab} = 0 = h$) the universal linearized equations of motion (3.9) now generalize to the following two equations for Einstein and curvature squared theories, respectively [50]:

$$\begin{aligned}\lambda_h \left(\bar{\square} + \frac{2}{\ell_{\text{AdS}}^2} \right) h_{ab} &= -\lambda_\phi T_{ab}^{(\phi)}, \\ \tilde{\lambda}_h \left(\bar{\square} + \frac{2}{\tilde{\ell}_{\text{AdS}}^2} \right) h_{ab} + \beta \ell^2 \left(\bar{\square} + \frac{2}{\tilde{\ell}_{\text{AdS}}^2} \right)^2 h_{ab} &= -\tilde{\lambda}_\phi \tilde{T}_{ab}^{(\phi)},\end{aligned}\quad (\text{A.4})$$

where the coefficients λ_h and $\tilde{\lambda}_h$ in (A.4) are now given as

$$\lambda_h = 1, \quad \tilde{\lambda}_h = \left(1 - 2(d(d+1)\alpha + d\beta + (d-2)(d-3)\gamma) \frac{\tilde{\ell}^2}{\tilde{\ell}_{\text{AdS}}^2} \right). \quad (\text{A.5})$$

We now want to argue that the second term in the second line of (A.4) can be neglected for our computation of entanglement entropy. Indeed, this term is obviously suppressed if β is treated as a small parameter and in the physical on-shell limit $(\bar{\square} + 2/\tilde{\ell}_{\text{AdS}}^2)h_{ab} \rightarrow 0$ (see also [61]). Further, we can also rewrite the second line of (A.4) in the following factorized form:

$$\beta \ell^2 \left(\bar{\square} + \frac{2}{\tilde{\ell}_{\text{AdS}}^2} \right) \left(\bar{\square} + \frac{2}{\tilde{\ell}_{\text{AdS}}^2} + \frac{\tilde{\lambda}_h}{\beta \ell^2} \right) h_{ab} = -\tilde{\lambda}_\phi T_{ab}^{(\phi)}. \quad (\text{A.6})$$

From this expression one can anticipate the two poles of the graviton propagator. The Fierz-Pauli type mass term $\sim 2/\tilde{\ell}_{\text{AdS}}^2$ corresponds to the physical graviton mode. The other mass term $\sim (2/\tilde{\ell}_{\text{AdS}}^2 + \tilde{\lambda}_h/\beta \ell^2)$ schematically leads to the ghost which is very heavy for small β and hence does not go on-shell at low energies. It can be seen as an artefact of truncating some well-behaved theory at four-derivative order.

A.2 General argument

Despite the previous subsection appearing to be restricted to the specific case of curvature squared theories, the conclusion is actually more general. If higher curvature terms, which appear in the gravitational action with some dimensionless coupling α , lead to ghosts modes, then the latter generically have a mass which scales with $1/\alpha$. If these modes are interpreted as due to consistently truncating perturbative quantum gravity (i.e., $|\alpha| \ll 1$), the ghosts are unphysical and indeed never go on-shell at low energies. In such a perturbative framework the presence of ghosts is hence not due to an instability of the theory (the underlying full theory is assumed to be unitary), but it is just a rather harmless consequence of neglecting certain sectors of the theory which are not being probed at low energies.

To illustrate these statements, consider, for example, a gravity sector of the form $\mathcal{L}_{\text{grav}} = \mathcal{L}_{\text{grav}}(g_{ab}, R_{abcd})$ such that the equations of motion are fourth order in derivatives. The most general four-derivative gravitational equations of motion linearized around AdS are precisely the same as those captured by the curvature squared theories in (A.1). This can be shown using the method of background field expansion; see, e.g., [67], where it was indeed argued that the linearized action of any higher derivative theory of this type is the same as that of the most general curvature squared theory.¹⁶ Concretely, given some such higher curvature theory, the action expanded to second order around AdS takes the form (A.1) with parameters $\{\alpha, \beta, \gamma\}$ determined as linear combinations of the original higher derivative couplings. We can therefore ignore the ghosts of these theories as long as the higher derivative couplings are viewed as a semiclassical truncation stemming from some unitary UV-complete theory.

B Perturbations of Wald entropy: background field expansion

This appendix complements the analysis of section 2.3.3 by computing the expansion of the Wald entropy density at $\mathcal{O}(h_{ab})$, derived from the general second order Lagrangian (2.21). This analysis will also serve to illuminate how the universality of graviton propagation on AdS simplifies the method of background field expansion, thus demonstrating a more general universality statement about quantities which depend on both a_d^* and C_T . Indeed, we will show in what sense Einstein-Gauss-Bonnet theory is the most general theory at $\mathcal{O}(h^2)$. Any other physically acceptable higher curvature theory has the same background field expansion at this order.

It is straightforward that we can focus on the pure gravity part of the general Lagrangian (2.21) since the minimally coupled matter sector does not couple to curvature. We wish to give a similar argument for the constant term. Ignoring the matter sector puts us in the realm of [68] (see also [6]), where a detailed analysis of perturbed Wald entropy on a maximally symmetric AdS background has been performed. The best strategy to compute the perturbative expansion of Wald entropy is to rewrite the expansion in h_{ab} as a background field expansion of curvatures R_{abcd} around $-\frac{1}{\ell_{\text{AdS}}^2}(g_{ac}g_{bd} - g_{ad}g_{bc})$. More explicitly, instead of writing explicit metric perturbations h_{ab} we write the Lagrangian solely

¹⁶In this context, see also section B.

in terms of

$$\begin{aligned}\Delta R_{abcd} &= R_{abcd} - \left[-\frac{1}{\ell_{\text{AdS}}^2} (g_{ac}g_{bd} - g_{ad}g_{bc}) \right] \\ &= {}^{(1)}R_{abcd} + \frac{1}{\ell_{\text{AdS}}^2} (\bar{g}_{ac}h_{bd} + \bar{g}_{bd}h_{ac} - \bar{g}_{ad}h_{bc} - \bar{g}_{bc}h_{ad}) + \mathcal{O}(h^2),\end{aligned}\tag{B.1}$$

where $g_{ab} = \bar{g}_{ab} + h_{ab}$ is the full metric and indices of expressions like this are raised with the full metric. This way, the pure gravity part of our general quadratic Lagrangian (2.21) can be interpreted as an expansion in ΔR_{abcd} and its contractions and derivatives. To wit, we can always write the pure gravitational couplings of (2.21) as a linear combination of all possible combinations of ΔR_{abcd} which contain terms relevant at $\mathcal{O}(h^2)$. Clearly such combinations contain at most two factors of ΔR_{abcd} , but they may contain any number of derivatives in general:

$$\begin{aligned}I_{\text{grav}} &= \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-\bar{g}} \left(-\frac{2d}{\ell_{\text{AdS}}^2} \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{\ell_{\text{P}}^{d-1}}{\ell_{\text{AdS}}^{d-1}} a_d^* + \frac{\lambda_h}{2} h_{ab} \mathcal{G}^{abcd} h_{cd} + \mathcal{O}(h^3) \right) \\ &= \frac{1}{2\ell_{\text{P}}^{d-1}} \int \sqrt{-g} \left(\mathcal{L}_{\Delta R}^{(I)} + \mathcal{L}_{\Delta R}^{(II)} \right),\end{aligned}\tag{B.2}$$

with

$$\begin{cases} \mathcal{L}_{\Delta R}^{(I)} \equiv a_0 + a_2 \Delta R \\ \mathcal{L}_{\Delta R}^{(II)} \equiv b_1 \ell_{\text{AdS}}^2 (\Delta R)^2 + b_2 \ell_{\text{AdS}}^2 \Delta R^{ab} \Delta R_{ab} + b_3 \ell_{\text{AdS}}^2 \Delta R^{abcd} \Delta R_{abcd} + \dots + \mathcal{O}(h^3) \end{cases}$$

Let us explain this way of parameterizing the action piece by piece. First observe that from the form of (B.2) it follows that a_0 is just

$$a_0 = -\frac{2d}{\ell_{\text{AdS}}^2} \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{\ell_{\text{P}}^{d-1}}{\ell_{\text{AdS}}^{d-1}} a_d^*.\tag{B.3}$$

Similarly, one can easily verify that the vanishing of the linearized Lagrangian (i.e. the equation of motion for the background solution \bar{g}_{ab}) demands $a_0 = -(2d/\ell_{\text{AdS}}^2)a_2$. This fixes the form of $\mathcal{L}_{\Delta R}^{(I)}$ and shows that the latter is all that would ever show up in the case of Einstein gravity (with an overall pre-factor a_2). Therefore, $\mathcal{L}_{\Delta R}^{(II)}$ is the term which parameterizes contributions to the graviton propagator which only come from higher curvature terms. More tensor structures involving derivatives and contractions of $\Delta R_{abcd}\Delta R_{efgh}$ could be written in (B.2) in order to match the expansion of an arbitrary theory of gravity at higher orders in h_{ab} . However, this is not needed to reproduce the first line, i.e., the universal parts up to $\mathcal{O}(h^2)$. To wit, the two lines of (B.2) match if we choose

$$b_1 = -\frac{b_2}{4} = b_3 = \frac{1}{4(d-2)} \left(\lambda_h - \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{\ell_{\text{P}}^{d-1}}{\ell_{\text{AdS}}^{d-1}} a_d^* \right).\tag{B.4}$$

Note that this choice is schematically of the Gauss-Bonnet form. We can interpret this as the statement that a Gauss-Bonnet term is all that is needed to reproduce the linearized graviton action with arbitrary graviton normalization. This is a generalization of the results of the main text: in CFT language, a quantity which is sensitive to both a_d^* and C_T

may undeniably detect the presence of higher curvature terms, but it does not distinguish between Gauss-Bonnet theory and any other higher curvature theory in the same way that quantities only knowing about a_d^* can always be computed using Einstein gravity. The reason is that linearized actions up to $\mathcal{O}(h^2)$ only require a small number of terms in the background field expansion.

As a consistency check, we can make the relation between parameters $\{a_0, a_2, b_1, b_2, b_3\}$ and C_T manifest by referring to the results of [68] (see also [6]). There it was shown that an action of the form (B.2) leads to $\lambda_h = a_2 + 4(d-2)b_3$ which is proportional to C_T via (2.33). This is in perfect agreement with the above statements.

Since (B.2) is manifestly written as both a perturbative expansion in the metric fluctuation and as a functional of curvatures, we can easily compute the associated perturbative expansion of Wald entropy density by taking derivatives with respect to R_{abcd} . Note that the way the expansion is written makes it explicit that any variation with respect to R_{abcd} reduces the order in h_{ab} by precisely one and the number of derivatives by two. The unperturbed Wald entropy is hence only sensitive to $\mathcal{L}_{\Delta R}^{(I)}$ and is readily checked to agree with (2.5):

$$S_{\text{Wald}}^{(0)}(\mathcal{E}_{\min}) = -2\pi \frac{1}{2\ell_{\text{P}}^{d-1}} \int_{\mathcal{E}_{\min}} \sqrt{\gamma} \left[\frac{\delta \mathcal{L}_{\Delta R}^{(I)}}{\delta R_{abcd}} n_{ab} n_{cd} \right]_{\text{AdS}} = \frac{2\pi}{\pi^{d/2}} \Gamma(d/2) a_d^* \frac{\int_{\mathcal{E}_{\min}} \sqrt{\gamma}}{\ell_{\text{AdS}}^{d-1}}, \quad (\text{B.5})$$

where the subscript ‘AdS’ refers to evaluation on the background \bar{g}_{ab} .

Similarly, the perturbation of Wald entropy is now computable via the two contributions

$$\delta S_{\text{Wald}}(\mathcal{E}_{\min}) = -2\pi \frac{1}{2\ell_{\text{P}}^{d-1}} \int_{\mathcal{E}_{\min}} \left[\sqrt{\gamma} \frac{\delta(\mathcal{L}_{\Delta R}^{(I)} + \mathcal{L}_{\Delta R}^{(II)})}{\delta R_{abcd}} n_{ab} n_{cd} \right]_{\mathcal{O}(h)}, \quad (\text{B.6})$$

where ‘ $\mathcal{O}(h)$ ’ picks out the term proportional to h_{ab} . We find the following contribution from $\mathcal{L}_{\Delta R}^{(I)}$:

$$\delta S_{\text{Wald}, (I)}(\mathcal{E}_{\min}) = -2\pi \frac{1}{2\ell_{\text{P}}^{d-1}} \int_{\mathcal{E}_{\min}} [\sqrt{\gamma} (-2a_2)]_{\mathcal{O}(h)} = \frac{2\pi}{\pi^{d/2}} \Gamma(d/2) a_d^* \frac{\int_{\mathcal{E}_{\min}} \sqrt{\gamma} (\frac{1}{2} \bar{\gamma}^{\alpha\beta} h_{\alpha\beta})}{\ell_{\text{AdS}}^{d-1}}, \quad (\text{B.7})$$

which could also be obtained by simply perturbing the volume form in (B.5). Similarly, one can compute the higher curvature contributions from $\mathcal{L}_{\Delta R}^{(II)}$ at $\mathcal{O}(h)$ as follows:¹⁷

$$\begin{aligned} \delta S_{\text{Wald}, (II)}(\mathcal{E}_{\min}) &= \\ &= -2\pi \frac{1}{2\ell_{\text{P}}^{d-1}} \int_{\mathcal{E}_{\min}} \sqrt{\gamma} 2b_1 \ell_{\text{AdS}}^2 \left[\Delta R g^{c[a} g^{b]d} - 4\Delta R_{ef} g^{e[b} g^{a][c} g^{d]f} + \Delta R^{abcd} \right] n_{ab} n_{cd} \\ &= 4\pi b_1 \ell_{\text{AdS}}^2 \frac{1}{\ell_{\text{P}}^{d-1}} \int_{\mathcal{E}_{\min}} \sqrt{\gamma} \left[R + 2R^{ab} n_{ac} n_b{}^c - \frac{1}{2} R^{abcd} n_{ab} n_{cd} \right]_{\mathcal{O}(h)} + \mathcal{O}(h^2) \\ &= 4\pi b_1 \ell_{\text{AdS}}^2 \frac{1}{\ell_{\text{P}}^{d-1}} \int_{\mathcal{E}_{\min}} \sqrt{\gamma} {}^{(1)}\mathcal{R} + \mathcal{O}(h^2), \end{aligned} \quad (\text{B.8})$$

¹⁷Our conventions are as follows: the embedded version of the intrinsic metric $\gamma_{\alpha\beta}$ is $\gamma_{ab} = g_{ab} + t_a t_b - n_a n_b$ where t^a and n^a are normal to the minimal surface such that $t^2 = -1$, $n^2 = 1$, $t^a n_a = 0$ and the binormal $n^{ab} \equiv n^a t^b - t^a n^b$.

where $^{(1)}\mathcal{R}$ is the linearized intrinsic Ricci scalar on \mathcal{E}_{\min} in units of ℓ_{AdS} (i.e., (B.8) is essentially the first order perturbation of Wald entropy in pure Gauss-Bonnet theory). In the last step we used the fact that extrinsic curvatures vanish on \mathcal{E}_{\min} , so the projected Ricci scalar in the penultimate line coincides with the intrinsic one. We can further simplify the expression using

$$^{(1)}\mathcal{R} = -\bar{\mathcal{R}}^{\alpha\beta} \delta\gamma_{\alpha\beta} + \bar{\nabla}_\alpha(\dots) = \frac{(d-2)}{\ell_{\text{AdS}}^2} \bar{\gamma}^{\alpha\beta} h_{\alpha\beta} + \bar{\nabla}_\alpha(\dots), \quad (\text{B.9})$$

where $\bar{\mathcal{R}}^{\alpha\beta}$ is the unperturbed intrinsic Ricci scalar and $\bar{\nabla}_\alpha$ is the intrinsic covariant derivative compatible with $\bar{\gamma}_{\alpha\beta}$ and hence only yields boundary terms which we are not keeping track of.¹⁸ Hence (B.8) becomes

$$\delta S_{\text{Wald, (II)}}(\mathcal{E}_{\min}) = 2\pi \left(\lambda_h - \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{\ell_{\text{P}}^{d-1}}{\ell_{\text{AdS}}^{d-1}} a_d^* \right) \frac{\int_{\mathcal{E}_{\min}} \sqrt{\bar{\gamma}} \left(\frac{1}{2} \bar{\gamma}^{\alpha\beta} h_{\alpha\beta} \right)}{\ell_{\text{P}}^{d-1}} \quad (\text{B.10})$$

Finally adding the two contributions to perturbed Wald entropy, (B.7) and (B.10), we conclude that (B.6) reads

$$\delta S_{\text{Wald}}(\mathcal{E}_{\min}) = 2\pi \left(\frac{\ell_{\text{AdS}}^{d-1}}{\ell_{\text{P}}^{d-1}} \lambda_h \right) \frac{\int_{\mathcal{E}_{\min}} \sqrt{\bar{\gamma}} \left(\frac{1}{2} \bar{\gamma}^{\alpha\beta} h_{\alpha\beta} \right)}{\ell_{\text{AdS}}^{d-1}}. \quad (\text{B.11})$$

C First law of entanglement

In the main text we have focused on perturbations of entanglement entropy due to scalar deformations. Let us now drop such particular assumptions and consider a slightly more general setup which has been argued to serve as a quantitative derivation of bulk dynamics from entanglement. For theories in which the Ryu-Takayanagi proposal is valid, a first law of entanglement entropy has been shown to be equivalent to linearized bulk equations of motion [4–6, 25, 27, 28].

The entanglement first law

$$\delta S_{EE}(\mathcal{A}) = \text{Tr}(\delta\rho_{\mathcal{A}} H_{\mathcal{A}}) \equiv \delta\langle H_{\mathcal{A}} \rangle \quad (\text{C.1})$$

is purely a CFT statement which says that any change of $S_{EE}(\mathcal{A})$ due to some small perturbation $\delta\rho_{\mathcal{A}}$ of the density matrix is given by the change of modular energy associated with the spatial sphere \mathcal{A} . (This statement can easily be derived from the formalism of section 2.2.) For CFTs with a semiclassical gravity dual, this statement has a natural analog in gravity: the left hand side of (C.1) corresponds to a perturbation of Wald entropy as in (2.34). Similarly the right hand side of (C.1) translates to a change of gravitational energy $E_{\text{grav}}(\mathcal{A})$ of the Rindler patch of AdS_{d+1} associated with the causal diamond of \mathcal{A} . That is, the CFT entanglement first law (C.1) is equivalent to the gravitational first law

$$\delta S_{\text{Wald}}(\mathcal{E}_{\min}) = \delta E_{\text{grav}}(\mathcal{A}). \quad (\text{C.2})$$

¹⁸Boundary terms are important to get a finite result as in (2.16). Since we are only interested in identifying the fate of the dimensionless parameters, we are being implicit about these subtleties and assume that there always exist suitable Gibbons-Hawking-York boundary terms.

The gravitational energy $E_{\text{grav}}(\mathcal{A})$ is defined as an integral of the holographic stress tensor T_{ab}^{grav} over a spatial bulk surface of bulk codimension two, which lies inside the Rindler patch of \mathcal{A} and ends on $\partial\mathcal{A}$. Without loss of generality, we push this bulk surface to the boundary, so the gravitational energy can be computed as a boundary integral over \mathcal{A} :

$$E_{\text{grav}}(\mathcal{A}) = \int_{\mathcal{A}} d\Sigma^\mu T_{\mu\nu}^{\text{grav}} \zeta_{\mathcal{A}}^\nu, \quad (\text{C.3})$$

where $d\Sigma^\mu$ is the volume form on the $(d-1)$ -dimensional surface \mathcal{A} and $\zeta_{\mathcal{A}}^\nu$ is the conformal Killing vector generating the Rindler flow inside the causal development of \mathcal{A} . In holographic CFTs the CFT stress tensor $T_{\mu\nu}$ of (2.13) agrees with the holographic stress tensor $T_{\mu\nu}^{\text{grav}}$. One can easily verify that the gravitational energy (C.3) then agrees with the definition of modular energy in the CFT, (2.12).

The holographic stress tensor is determined by the dominant asymptotics of the graviton fluctuation on AdS_{d+1} . Let us work in Poincaré coordinates

$$ds^2 = \frac{\ell_{\text{AdS}}^2}{z^2} (dz^2 + \eta_{\mu\nu} dx^\mu dx^\nu) + h_{ab} dx^a dx^b. \quad (\text{C.4})$$

Then the relevant graviton falloff near the boundary is given by the following piece of the Fefferman-Graham expansion:

$$h_{ab}(x^\rho, z) dx^a dx^b = \ell_{\text{AdS}}^2 z^{d-2} h_{\mu\nu}^{(d)}(x^\rho) dx^\mu dx^\nu + \dots \quad (\text{C.5})$$

The power of z is chosen such that the induced boundary stress tensor $\delta T_{\mu\nu}^{\text{grav}}$ is finite (i.e., z -independent). It was shown in [6, 68] that for a wide class of theories the holographic stress tensor associated with this bulk perturbation evaluates to

$$\delta T_{\mu\nu}^{\text{grav}} = \frac{d C_T}{f_d} h_{\mu\nu}^{(d)}, \quad (\text{C.6})$$

where f_d is the number defined in (2.33). Note that this form of $\delta T_{\mu\nu}^{\text{grav}}$ is consistent with our considerations in section 2.3.3. Indeed, $\delta S_{\text{Wald}}(\mathcal{E}_{\text{min}})$ in the present context can be obtained from the gravitational energy integral over the combination (C.6):

$$\delta S_{\text{Wald}}(\mathcal{E}_{\text{min}}) = \frac{d C_T}{f_d} \int_{\mathcal{A}} d\Sigma^\mu h_{\mu\nu}^{(d)} \zeta_{\mathcal{A}}^\nu. \quad (\text{C.7})$$

This has the same structure as (2.36), i.e., an integral over a quantity (C_T or λ_h times a metric perturbation) which satisfies a universal equation of motion.

The linearized equations of motion can be derived from the first law statements above. Indeed, assuming the Ryu-Takayanagi conjecture, it was shown in [6, 68] that (C.2) for small spherical regions \mathcal{A} is equivalent to linearized gravitational equations of motion for any higher derivative Lagrangian $\mathcal{L}_{\text{grav}} = \mathcal{L}_{\text{grav}}(g_{ab}, R_{abcd}, \nabla_e R_{abcd}, \dots)$. As argued in section 2.3, on an AdS background these equations are the same for any theory of gravity with second order equations of motion, up to overall normalizations which encode the higher curvature contributions. Depending on the precise setup, even these normalizations may be absorbed into renormalizations of dimensionless quantities. In particular the normalization λ_h inside C_T always multiplies $(\ell_{\text{AdS}}/\ell_{\text{P}})^{d-1}$, cf., (2.33). Unless $\ell_{\text{AdS}}/\ell_{\text{P}}$ has already been fixed (e.g., by a condition such as (2.7)), this ratio can serve to account for any λ_h dependence.

D Second order gravitational perturbation theory

In this appendix we list some relevant formulae for gravitational perturbation theory up to second order in the graviton. We take the ansatz

$$g_{ab} = \bar{g}_{ab} + h_{ab} \quad (\text{D.1})$$

and are interested in perturbations up to $\mathcal{O}(h^2)$. The background solution \bar{g}_{ab} is AdS_{d+1} with curvature $-1/\ell_{\text{AdS}}^2$, so the following simplifications apply for the background curvature tensors:

$$\bar{R}_{abcd} = -\frac{1}{\ell_{\text{AdS}}^2}(\bar{g}_{ac}\bar{g}_{bd} - \bar{g}_{ad}\bar{g}_{bc}), \quad \bar{R}_{ab} = -\frac{d}{\ell_{\text{AdS}}^2}\bar{g}_{ab}, \quad \bar{R} = -\frac{d(d+1)}{\ell_{\text{AdS}}^2}. \quad (\text{D.2})$$

Let us evaluate and define various quantities, where superscript numbers refer to the order in perturbation theory:

$$\begin{aligned} g^{ab} &= \bar{g}^{ab} - h^{ab} + h^{ac}h_c^b + \mathcal{O}(h^3), \\ \Gamma_{ab}^\rho &= \bar{\Gamma}_{ab}^\rho + {}^{(1)}\Gamma_{ab}^c - h_d^c {}^{(1)}\Gamma_{ab}^d + \mathcal{O}(h^3), \\ R^a{}_{bcd} &= \bar{R}^a{}_{bcd} + {}^{(1)}R^a{}_{bcd} \\ &\quad - \left[h_e^a {}^{(1)}R^e{}_{bcd} + \bar{g}^{ae}\bar{g}_{fg} \left({}^{(1)}\Gamma_{ce}^g {}^{(1)}\Gamma_{db}^f - {}^{(1)}\Gamma_{de}^g {}^{(1)}\Gamma_{cb}^f \right) \right] + \mathcal{O}(h^3), \\ R_{ab} &= \bar{R}_{ab} + {}^{(1)}R_{ab} - \left[h_d^c {}^{(1)}R^d{}_{acb} + \bar{g}^{ce}\bar{g}_{fg} \left({}^{(1)}\Gamma_{ce}^g {}^{(1)}\Gamma_{ba}^f - {}^{(1)}\Gamma_{ce}^g {}^{(1)}\Gamma_{ca}^f \right) \right] + \mathcal{O}(h^3), \\ R &= \bar{R} + {}^{(1)}R + \left\{ h^{ac}h_c^b \bar{R}_{ab} - h^{ab} {}^{(1)}R_{ab} \right. \\ &\quad \left. - \bar{g}^{ab} \left[h_d^c {}^{(1)}R^d{}_{acb} + \bar{g}^{ce}\bar{g}_{fg} \left({}^{(1)}\Gamma_{ce}^g {}^{(1)}\Gamma_{ba}^f - {}^{(1)}\Gamma_{be}^g {}^{(1)}\Gamma_{ca}^f \right) \right] \right\} + \mathcal{O}(h^3). \end{aligned} \quad (\text{D.3})$$

In these expressions the linearized objects are given in terms of h_{ab} by

$$\begin{aligned} {}^{(1)}\Gamma_{ab}^c &\equiv \frac{1}{2}\bar{g}^{cd}(\bar{\nabla}_a h_{bd} + \bar{\nabla}_b h_{ad} - \bar{\nabla}_d h_{ab}), \\ {}^{(1)}R^a{}_{bcd} &= \bar{\nabla}_{[c}\bar{\nabla}_{d]}h_b^a + \bar{\nabla}_{[c}\bar{\nabla}_b h_{d]}^a - \bar{\nabla}_{[c}\bar{\nabla}^a h_{d]}b, \\ {}^{(1)}R_{ab} &= \bar{\nabla}_c \bar{\nabla}_{(a} h_{b)}^c - \frac{1}{2}\bar{\square}h_{ab} - \frac{1}{2}\bar{\nabla}_a \bar{\nabla}_b h, \\ {}^{(1)}R &= {}^{(1)}R_{ab}\bar{g}^{ab} - \bar{R}^{ab}h_{ab} = \bar{\nabla}_a \bar{\nabla}_b h^{ab} - \bar{\square}h - \bar{R}^{ab}h_{ab}, \end{aligned} \quad (\text{D.4})$$

where $h \equiv h_a^a = \bar{g}^{ab}h_{ab}$.

For the expansion of actions it is also useful to note the following identity:

$$\sqrt{-g} = \sqrt{-\bar{g}} \left(1 + \frac{1}{2}h + \frac{1}{8}(h^2 - 2h_b^a h_a^b) + \mathcal{O}(h^3) \right). \quad (\text{D.5})$$

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] O. Aharony, S.S. Gubser, J.M. Maldacena, H. Ooguri and Y. Oz, *Large- N field theories, string theory and gravity*, *Phys. Rept.* **323** (2000) 183 [[hep-th/9905111](#)] [[INSPIRE](#)].
- [2] M. Van Raamsdonk, *Comments on quantum gravity and entanglement*, [arXiv:0907.2939](#) [[INSPIRE](#)].
- [3] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [[arXiv:1005.3035](#)] [[INSPIRE](#)].
- [4] N. Lashkari, M.B. McDermott and M. Van Raamsdonk, *Gravitational dynamics from entanglement ‘thermodynamics’*, *JHEP* **04** (2014) 195 [[arXiv:1308.3716](#)] [[INSPIRE](#)].
- [5] D.D. Blanco, H. Casini, L.-Y. Hung and R.C. Myers, *Relative entropy and holography*, *JHEP* **08** (2013) 060 [[arXiv:1305.3182](#)] [[INSPIRE](#)].
- [6] T. Faulkner, M. Guica, T. Hartman, R.C. Myers and M. Van Raamsdonk, *Gravitation from entanglement in holographic CFTs*, *JHEP* **03** (2014) 051 [[arXiv:1312.7856](#)] [[INSPIRE](#)].
- [7] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [[hep-th/0603001](#)] [[INSPIRE](#)].
- [8] S. Ryu and T. Takayanagi, *Aspects of holographic entanglement entropy*, *JHEP* **08** (2006) 045 [[hep-th/0605073](#)] [[INSPIRE](#)].
- [9] V.E. Hubeny, M. Rangamani and T. Takayanagi, *A covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062 [[arXiv:0705.0016](#)] [[INSPIRE](#)].
- [10] H. Casini, M. Huerta and R.C. Myers, *Towards a derivation of holographic entanglement entropy*, *JHEP* **05** (2011) 036 [[arXiv:1102.0440](#)] [[INSPIRE](#)].
- [11] T. Faulkner, *Bulk emergence and the RG flow of entanglement entropy*, *JHEP* **05** (2015) 033 [[arXiv:1412.5648](#)] [[INSPIRE](#)].
- [12] S. de Haro, S.N. Solodukhin and K. Skenderis, *Holographic reconstruction of space-time and renormalization in the AdS/CFT correspondence*, *Commun. Math. Phys.* **217** (2001) 595 [[hep-th/0002230](#)] [[INSPIRE](#)].
- [13] L.-Y. Hung, R.C. Myers and M. Smolkin, *Some calculable contributions to holographic entanglement entropy*, *JHEP* **08** (2011) 039 [[arXiv:1105.6055](#)] [[INSPIRE](#)].
- [14] H. Liu and M. Mezei, *A refinement of entanglement entropy and the number of degrees of freedom*, *JHEP* **04** (2013) 162 [[arXiv:1202.2070](#)] [[INSPIRE](#)].
- [15] T. Nishioka, *Relevant perturbation of entanglement entropy and stationarity*, *Phys. Rev. D* **90** (2014) 045006 [[arXiv:1405.3650](#)] [[INSPIRE](#)].
- [16] V. Rosenhaus and M. Smolkin, *Entanglement entropy: a perturbative calculation*, *JHEP* **12** (2014) 179 [[arXiv:1403.3733](#)] [[INSPIRE](#)].
- [17] V. Rosenhaus and M. Smolkin, *Entanglement entropy for relevant and geometric perturbations*, *JHEP* **02** (2015) 015 [[arXiv:1410.6530](#)] [[INSPIRE](#)].
- [18] D.V. Fursaev, A. Patrushev and S.N. Solodukhin, *Distributional geometry of squashed cones*, *Phys. Rev. D* **88** (2013) 044054 [[arXiv:1306.4000](#)] [[INSPIRE](#)].
- [19] L.-Y. Hung, R.C. Myers and M. Smolkin, *On holographic entanglement entropy and higher curvature gravity*, *JHEP* **04** (2011) 025 [[arXiv:1101.5813](#)] [[INSPIRE](#)].

- [20] J. de Boer, M. Kulaxizi and A. Parnachev, *Holographic entanglement entropy in Lovelock gravities*, *JHEP* **07** (2011) 109 [[arXiv:1101.5781](#)] [[INSPIRE](#)].
- [21] X. Dong, *Holographic entanglement entropy for general higher derivative gravity*, *JHEP* **01** (2014) 044 [[arXiv:1310.5713](#)] [[INSPIRE](#)].
- [22] J. Camps, *Generalized entropy and higher derivative gravity*, *JHEP* **03** (2014) 070 [[arXiv:1310.6659](#)] [[INSPIRE](#)].
- [23] A. Bhattacharyya, M. Sharma and A. Sinha, *On generalized gravitational entropy, squashed cones and holography*, *JHEP* **01** (2014) 021 [[arXiv:1308.5748](#)] [[INSPIRE](#)].
- [24] A. Bhattacharyya and M. Sharma, *On entanglement entropy functionals in higher derivative gravity theories*, *JHEP* **10** (2014) 130 [[arXiv:1405.3511](#)] [[INSPIRE](#)].
- [25] B. Swingle and M. Van Raamsdonk, *Universality of gravity from entanglement*, [arXiv:1405.2933](#) [[INSPIRE](#)].
- [26] J. Bhattacharya, M. Nozaki, T. Takayanagi and T. Ugajin, *Thermodynamical property of entanglement entropy for excited states*, *Phys. Rev. Lett.* **110** (2013) 091602 [[arXiv:1212.1164](#)] [[INSPIRE](#)].
- [27] G. Wong, I. Klich, L.A. Pando Zayas and D. Vaman, *Entanglement temperature and entanglement entropy of excited states*, *JHEP* **12** (2013) 020 [[arXiv:1305.3291](#)] [[INSPIRE](#)].
- [28] D. Allahbakhshi, M. Alishahiha and A. Naseh, *Entanglement thermodynamics*, *JHEP* **08** (2013) 102 [[arXiv:1305.2728](#)] [[INSPIRE](#)].
- [29] V. Iyer and R.M. Wald, *Some properties of Noether charge and a proposal for dynamical black hole entropy*, *Phys. Rev. D* **50** (1994) 846 [[gr-qc/9403028](#)] [[INSPIRE](#)].
- [30] R.C. Myers and A. Sinha, *Holographic c-theorems in arbitrary dimensions*, *JHEP* **01** (2011) 125 [[arXiv:1011.5819](#)] [[INSPIRE](#)].
- [31] R.C. Myers and A. Sinha, *Seeing a c-theorem with holography*, *Phys. Rev. D* **82** (2010) 046006 [[arXiv:1006.1263](#)] [[INSPIRE](#)].
- [32] A. Schwimmer and S. Theisen, *Entanglement entropy, trace anomalies and holography*, *Nucl. Phys. B* **801** (2008) 1 [[arXiv:0802.1017](#)] [[INSPIRE](#)].
- [33] J. Erdmenger and H. Osborn, *Conserved currents and the energy momentum tensor in conformally invariant theories for general dimensions*, *Nucl. Phys. B* **483** (1997) 431 [[hep-th/9605009](#)] [[INSPIRE](#)].
- [34] H. Osborn and A.C. Petkou, *Implications of conformal invariance in field theories for general dimensions*, *Annals Phys.* **231** (1994) 311 [[hep-th/9307010](#)] [[INSPIRE](#)].
- [35] V. Ogievetsky and I. Polubarinov, *Interacting field of spin 2 and the Einstein equations*, *Annals Phys.* **35** (1965) 167.
- [36] R.M. Wald, *Spin-2 fields and general covariance*, *Phys. Rev. D* **33** (1986) 3613 [[INSPIRE](#)].
- [37] K.S. Stelle, *Classical gravity with higher derivatives*, *Gen. Rel. Grav.* **9** (1978) 353 [[INSPIRE](#)].
- [38] K.S. Stelle, *Renormalization of higher derivative quantum gravity*, *Phys. Rev. D* **16** (1977) 953 [[INSPIRE](#)].
- [39] N.H. Barth and S.M. Christensen, *Quantizing fourth order gravity theories. 1. The functional integral*, *Phys. Rev. D* **28** (1983) 1876 [[INSPIRE](#)].

- [40] T. Biswas, E. Gerwick, T. Koivisto and A. Mazumdar, *Towards singularity and ghost free theories of gravity*, *Phys. Rev. Lett.* **108** (2012) 031101 [[arXiv:1110.5249](#)] [[INSPIRE](#)].
- [41] A. Conroy, A. Mazumdar and A. Teimouri, *Wald entropy for ghost-free, infinite derivative theories of gravity*, *Phys. Rev. Lett.* **114** (2015) 201101 [[arXiv:1503.05568](#)] [[INSPIRE](#)].
- [42] T.P. Sotiriou and V. Faraoni, *$f(R)$ theories of gravity*, *Rev. Mod. Phys.* **82** (2010) 451 [[arXiv:0805.1726](#)] [[INSPIRE](#)].
- [43] D.G. Boulware, S. Deser and K. Stelle, *Properties of energy in higher derivative gravity theories*, in *Quantum field theory and quantum statistics*, vol. 2, I.A. Batalin et al. eds., U.S.A. (1987), pg. 101 [[INSPIRE](#)].
- [44] E. Alvarez, D. Blas, J. Garriga and E. Verdaguer, *Transverse Fierz-Pauli symmetry*, *Nucl. Phys. B* **756** (2006) 148 [[hep-th/0606019](#)] [[INSPIRE](#)].
- [45] P. Van Nieuwenhuizen, *On ghost-free tensor Lagrangians and linearized gravitation*, *Nucl. Phys. B* **60** (1973) 478 [[INSPIRE](#)].
- [46] D. Lovelock, *The Einstein tensor and its generalizations*, *J. Math. Phys.* **12** (1971) 498 [[INSPIRE](#)].
- [47] R.C. Myers and B. Robinson, *Black holes in quasi-topological gravity*, *JHEP* **08** (2010) 067 [[arXiv:1003.5357](#)] [[INSPIRE](#)].
- [48] A. Buchel, J. Escobedo, R.C. Myers, M.F. Paulos, A. Sinha and M. Smolkin, *Holographic GB gravity in arbitrary dimensions*, *JHEP* **03** (2010) 111 [[arXiv:0911.4257](#)] [[INSPIRE](#)].
- [49] D.G. Boulware and S. Deser, *String generated gravity models*, *Phys. Rev. Lett.* **55** (1985) 2656 [[INSPIRE](#)].
- [50] S. Deser and B. Tekin, *Energy in generic higher curvature gravity theories*, *Phys. Rev. D* **67** (2003) 084009 [[hep-th/0212292](#)] [[INSPIRE](#)].
- [51] W.R. Kelly, K. Kuns and D. Marolf, *'t Hooft suppression and holographic entropy*, *JHEP* **10** (2015) 059 [[arXiv:1507.03654](#)] [[INSPIRE](#)].
- [52] M. Henningson and K. Skenderis, *The holographic Weyl anomaly*, *JHEP* **07** (1998) 023 [[hep-th/9806087](#)] [[INSPIRE](#)].
- [53] M. Blau, K.S. Narain and E. Gava, *On subleading contributions to the AdS/CFT trace anomaly*, *JHEP* **09** (1999) 018 [[hep-th/9904179](#)] [[INSPIRE](#)].
- [54] S. Nojiri and S.D. Odintsov, *On the conformal anomaly from higher derivative gravity in AdS/CFT correspondence*, *Int. J. Mod. Phys. A* **15** (2000) 413 [[hep-th/9903033](#)] [[INSPIRE](#)].
- [55] S.N. Solodukhin, *Entanglement entropy, conformal invariance and extrinsic geometry*, *Phys. Lett. B* **665** (2008) 305 [[arXiv:0802.3117](#)] [[INSPIRE](#)].
- [56] T. Hirata and T. Takayanagi, *AdS/CFT and strong subadditivity of entanglement entropy*, *JHEP* **02** (2007) 042 [[hep-th/0608213](#)] [[INSPIRE](#)].
- [57] E. Fradkin and J.E. Moore, *Entanglement entropy of 2D conformal quantum critical points: hearing the shape of a quantum drum*, *Phys. Rev. Lett.* **97** (2006) 050404 [[cond-mat/0605683](#)] [[INSPIRE](#)].
- [58] H. Casini and M. Huerta, *Universal terms for the entanglement entropy in 2 + 1 dimensions*, *Nucl. Phys. B* **764** (2007) 183 [[hep-th/0606256](#)] [[INSPIRE](#)].

- [59] R.C. Myers and A. Singh, *Entanglement entropy for singular surfaces*, *JHEP* **09** (2012) 013 [[arXiv:1206.5225](#)] [[INSPIRE](#)].
- [60] P. Bueno, R.C. Myers and W. Witczak-Krempa, *Universality of corner entanglement in conformal field theories*, *Phys. Rev. Lett.* **115** (2015) 021602 [[arXiv:1505.04804](#)] [[INSPIRE](#)].
- [61] P. Bueno and R.C. Myers, *Corner contributions to holographic entanglement entropy*, *JHEP* **08** (2015) 068 [[arXiv:1505.07842](#)] [[INSPIRE](#)].
- [62] M. Alishahiha, A.F. Astaneh, P. Fonda and F. Omidi, *Entanglement entropy for singular surfaces in hyperscaling violating theories*, *JHEP* **09** (2015) 172 [[arXiv:1507.05897](#)] [[INSPIRE](#)].
- [63] R.-X. Miao, *A holographic proof of the universality of corner entanglement for CFTs*, *JHEP* **10** (2015) 038 [[arXiv:1507.06283](#)] [[INSPIRE](#)].
- [64] L.-Y. Hung, R.C. Myers, M. Smolkin and A. Yale, *Holographic calculations of Rényi entropy*, *JHEP* **12** (2011) 047 [[arXiv:1110.1084](#)] [[INSPIRE](#)].
- [65] D.A. Galante and R.C. Myers, *Holographic Rényi entropies at finite coupling*, *JHEP* **08** (2013) 063 [[arXiv:1305.7191](#)] [[INSPIRE](#)].
- [66] E. Perlmutter, *A universal feature of CFT Rényi entropy*, *JHEP* **03** (2014) 117 [[arXiv:1308.1083](#)] [[INSPIRE](#)].
- [67] A.J. Amsel and D. Gorbos, *Wald-like formula for energy*, *Phys. Rev. D* **87** (2013) 024032 [[arXiv:1209.1603](#)] [[INSPIRE](#)].
- [68] K. Sen and A. Sinha, *Holographic stress tensor at finite coupling*, *JHEP* **07** (2014) 098 [[arXiv:1405.7862](#)] [[INSPIRE](#)].